On the Hardness of Approximating Minimum Vertex Cover

Irit Dinur* Samuel Safra†

May 26, 2004

Abstract

We prove the Minimum Vertex Cover problem to be NP-hard to approximate to within a factor of 1.3606, extending on previous PCP and hardness of approximation technique. To that end, one needs to develop a new proof framework, and borrow and extend ideas from several fields.

1 Introduction

The basic purpose of Computational Complexity Theory is to classify computational problems according to the amount of resources required to solve them. In particular, the most basic task is to classify computational problems to those that are efficiently solvable and those that are not. The complexity class P consists of all problems that can be solved in polynomial-time. It is considered, for this rough classification, as the class of efficiently-solvable problems. While many computational problems are known to be in P, many others, are neither known to be in P, nor proven to be outside P. Indeed many such problems are known to be in the class NP, namely the class of all problems whose solutions can be *verified* in polynomial-time. When it comes to proving that a problem is outside a certain complexity class, current techniques are radically inadequate. The most fundamental open question of Complexity Theory, namely, the P vs. NP question, may be a particular instance of this shortcoming.

While the P vs NP question is wide open, one may still classify computational problems into those in P and those that are *NP-hard* [Coo71, Lev73, Kar72]. A computational problem L is NP-hard if its complexity epitomizes the hardness of NP. That is, any NP problem can be *efficiently reduced* to L. Thus, the existence of a polynomial-time solution for L implies P=NP. Consequently, showing P$\neq$NP would immediately rule out an efficient algorithm for any NP-hard problem. Therefore, unless one intends to show NP=P, one should avoid trying to come up with an efficient algorithm for an NP-hard problem.

Let us turn our attention to a particular type of computational problems, namely, *optimization problems* — where one looks for an *optimal* among all plausible solutions. Some optimization problems are known to be NP-hard, for example, finding a largest size independent set in a graph [Coo71, Kar72], or finding an assignment satisfying the maximum number of clauses in a given 3CNF formula (MAX3SAT) [Kar72].

* The Miller Institute, UC Berkeley. Email: iritd@cs.berkeley.edu.

† School of Mathematics and School of Computer Science, Tel Aviv University and The Miller Institute, UC Berkeley. Research supported in part by the Fund for Basic Research administered by the Israel Academy of Sciences, and a Binational US-Israeli BSF grant. Email: safra@math.tau.ac.il.

A proof that some optimization problem is NP-hard, serves as an indication that one should relax the specification. A natural manner by which to do so is to require only an approximate solution — one that is not optimal, but is within a small factor $C > 1$ of optimal. Distinct optimization problems may differ significantly with regards to the optimal (closest to 1) factor C_{opt} to within which they can be efficiently approximated. Even optimization problems that are closely related, may turn out to be quite distinct with respect to C_{opt}. Let the **Maximum Independent Set** be the problem of finding, in a given graph G, the largest set of vertices that induces no edges. Let the **Minimum Vertex Cover** be the problem of finding the complement of this set (i.e. the smallest set of vertices that touch all edges). Clearly, for every graph G, a solution to Minimum Vertex Cover is (the complement of) a solution to Maximum Independent Set. However, the approximation behavior of these two problems is very different, as for Minimum Vertex Cover the value of C_{opt} is at most 2 [Hal02, BYE85, MS83], while for Maximum Independent Set it is at least $n^{1-\epsilon}$ [Hås99]. Classifying approximation problems according to their *approximation complexity* —namely, according to the optimal (closest to 1) factor C_{opt} to within which they can be efficiently approximated— has been investigated widely. A large body of work has been devoted to finding efficient approximation algorithms for a variety of optimization problems. Some NP-hard problems admit a polynomial-time approximation scheme (PTAS), which means they can be approximated, in polynomial time, to within any constant close to 1 (but not 1). Papadimitriou and Yannakakis [PY91] identified the class APX of problems (which includes for example Minimum Vertex Cover, Maximum Cut, and many others) and showed that either all problems in APX are NP-hard to approximate to within some factor bounded away from 1, or they all admit a PTAS.

The major turning point in the theory of approximability, was the discovery of the PCP Theorem [AS98, ALM$^+$98] and its connection to inapproximability [FGL$^+$96]. The PCP theorem immediately implies that all problems in APX are hard to approximate to within some constant factor. Much effort has been directed since towards a better understanding of the PCP methodology, thereby coming up with stronger and more refined characterizations of the class NP [AS98, ALM$^+$98, BGLR93, RS97, Hås99, Hås01]. The value of C_{opt} has been further studied (and in many cases essentially determined) for many classical approximation problems, in a large body of hardness-of-approximation results. For example, computational problems regarding lattices, were shown NP-hard to approximate [ABSS97, Ajt98, Mic, DKRS03] (to within factors still quite far from those achieved by the lattice basis reduction algorithm [LLL82]). Numerous combinatorial optimization problems were shown NP-hard to approximate to within a factor even marginally better than the best known efficient algorithm [LY94, BGS98, Fei98, FK98, Hås01, Hås99]. The approximation complexity of a handful of classical optimization problems is still open, namely, for these problems, the known upper and lower bounds for C_{opt} do not match.

One of these problems, and maybe the one that underscores the limitations of known technique for proving hardness of approximation, is Minimum Vertex Cover. Proving hardness for approximating Minimum Vertex Cover translates to obtaining a reduction of the following form. Begin with some NP-complete language L, and translate 'yes' instances $x \in L$ to graphs in which the largest independent set consists of a large fraction (up to half) of the vertices. 'No' instances $x \notin L$ translate to graphs in which the largest independent set is much smaller. Previous techniques resulted in graphs in which the ratio between the maximal independent set in the 'yes' and 'no' cases is very large (even $|V|^{1-\epsilon}$) [Hås99]. However, the maximal independent set in both 'yes' and 'no' cases, was very small $|V|^c$, for some $c < 1$. Håstad's celebrated paper [Hås01] achieving optimal inapproximability results in particular for linear equations mod 2, directly implies an inapproximability result for Minimum Vertex Cover of $\frac{7}{6}$. In this paper we go beyond that factor, proving the following theorem:

Theorem 1.1 *Given a graph G, it is NP-hard to approximate the minimum Vertex Cover to within any factor smaller than* $10\sqrt{5} - 21 = 1.3606\ldots$*.*

The proof proceeds by reduction, transforming instances of some NP-complete language L into graphs. We will (easily) prove that every 'yes'-instance (i.e. an input $x \in L$) is transformed into a graph that has a large independent set. The more interesting part will be to prove that every 'no'-instance (i.e. an input $x \notin L$) is transformed into a graph whose largest independent set is relatively small.

As it turns out, to that end, one has to apply several techniques and methods, stemming from distinct, seemingly unrelated, fields. Our proof incorporates theorems and insights from harmonic analysis of Boolean functions, and extremal set theory. Techniques which seem to be of independent interest. They have already shown application in proving hardness of approximation [DGKR03, DRS02, KR03], and would hopefully come handy in other areas.

Let us proceed to describe these techniques and how they relate to our construction. For the exposition, let us narrow the discussion and describe how to analyze independent sets in one specific graph, called the *non-intersection* graph. This graph is a key building-block in our construction. The formal definition of the non-intersection graph $G[n]$ is simple. Denote $[n] = \{1, ..., n\}$.

Definition 1.1 (Non-intersection graph) *$G[n]$ has one vertex for every subset $S \subseteq [n]$, and two vertices S_1 and S_2 are adjacent iff $S_1 \cap S_2 = \phi$.*

The final graph resulting from our reduction will be made of copies of $G[n]$ that are further inter-connected. Clearly, an independent set in the final graph is an independent set in each individual copy of $G[n]$.

To analyze our reduction, it is worthwhile to first analyze large independent sets in $G[n]$. It is useful to simultaneously keep in mind several equivalent perspectives of a set of vertices of $G[n]$, namely:

- A subset of the 2^n vertices of $G[n]$.
- A family of subsets of $[n]$.
- A Boolean function $\mathsf{f} : \{-1, 1\}^n \to \{-1, 1\}$. (Assign to every subset an n-bit string σ, with -1 in coordinates in the subset and 1 otherwise. Let $f(\sigma)$ be -1 or 1 depending on whether the subset is in the family or out).

In the remaining part of the introduction, we survey results from various fields on which we base our analysis. We first discuss issues related to Analysis of Boolean Functions, move on to describe some specific codes, and then discuss relevant issues in Extremal Set Theory. We end by describing the central feature of the new PCP construction, on which our entire approach hinges.

1.1 Analysis of Boolean Functions

Analysis of Boolean Functions can be viewed as harmonic analysis over the group $\mathbb{Z}_2^n$. Here tools from classical Harmonic Analysis are combined with techniques specific to functions of finite discrete range. Applications range from Social Choice, Economics and Game Theory, Percolation and Statistical Mechanics, and Circuit Complexity. This study has been carried out in recent years [BOL89, KKL88, BK97, FK96, BKS99], one of the outcomes of which is a theorem of Friedgut [Fri98] whose proof is based on the techniques introduced in [KKL88], which the proof herein utilizes in a critical manner. Let us briefly survey the fundamental principles of this field and the manner in which it is utilized.

Consider the group $\mathbb{Z}_2^n$. It will be convenient to view group elements as vectors in $\{-1,1\}^n$ with coordinate-wise multiplication as the group operation. Let f be a real-valued function on that group

$$\mathsf{f}\colon \{-1,1\}^n \to \mathbb{R}.$$

It is useful to view f as a vector in $\mathbb{R}^{2^n}$. We endow this space with an inner-product, $\mathsf{f}\cdot\mathsf{g} \stackrel{def}{=} \mathbf{E}_x\left[\mathsf{f}(x)\cdot\mathsf{g}(x)\right] = \frac{1}{2^n}\sum_x \mathsf{f}(x)\mathsf{g}(x)$. We associate each character of $\mathbb{Z}_2^n$ with a subset $S \subseteq [n]$ as follows,

$$\chi_S : \{-1,1\}^n \to \mathbb{R}, \qquad \chi_S(x) = \prod_{i\in S} x_i\,.$$

The set of characters $\{\chi_S\}_S$ forms an orthonormal basis for $\mathbb{R}^{2^n}$. The expansion of a function f in that basis is its Fourier-Walsh transform. The coefficient of χ_S in this expansion is denoted $\widehat{\mathsf{f}}(S) = \mathbf{E}_x\left[\mathsf{f}(x)\cdot\chi_S(x)\right]$, hence

$$\mathsf{f} = \sum_S \widehat{\mathsf{f}}(S)\cdot\chi_S\,.$$

Consider now the special case of a Boolean function f over the same domain

$$\mathsf{f}\colon \{-1,1\}^n \to \{-1,1\}.$$

Many natural operators and parameters of such an f have a neat and helpful formulation in terms of the Fourier-Walsh transform. This has yielded some striking results regarding Voting-Systems, Sharp-Threshold Phenomena, Percolation, and Complexity Theory.

The *influence* of a variable $i \in [n]$ on f is the probability, over a random choice of $x \in \{-1,1\}^n$, that flipping x_i changes the value of f:

$$\mathsf{influence}_i(\mathsf{f}) \stackrel{def}{=} \Pr\left[\mathsf{f}(x) \neq \mathsf{f}(x \odot \{i\})\right]$$

where $\{i\}$ is interpreted to be the vector that equals 1 everywhere except at the i'th coordinate where it equals -1, and $\odot$ denotes the group's multiplication.

The influence of the ith variable can be easily shown [BOL89] to be expressible in term of the Fourier coefficients of f as

$$\mathsf{influence}_i(\mathsf{f}) = \sum_{S\ni i} \widehat{\mathsf{f}}^2(S)\,.$$

The *total-influence* or *average-sensitivity* of f is the sum of influences

$$as(\mathsf{f}) \stackrel{def}{=} \sum_i \mathsf{influence}_i(\mathsf{f}) = \sum_S \widehat{\mathsf{f}}^2(S)\cdot|S|\,.$$

These notions (and others) regarding functions may also be examined for a non-uniform distribution over $\{-1,1\}^n$, in particular, for $0 < p < 1$, the *p-biased product-distribution* is

$$\mu_p(x) = p^{|x|}(1-p)^{n-|x|}$$

where $|x|$ is the number of -1's in x. One can define influence and average-sensitivity under the μ_p distribution, in much the same way. We have a different orthonormal basis for these functions [Tal94] because changing distributions changes the value of the inner-product of two functions.

Let $\mu_p(\mathsf{f})$ denote the probability that a given Boolean function f is -1. It is not hard to see that for monotone f, $\mu_p(\mathsf{f})$ increases with p. Moreover, the well-known Russo's Lemma [Mar74, Rus82] (Theorem 3.4) states that, for a monotone Boolean function f, the derivative $\frac{d\mu_p(\mathsf{f})}{dp}$ (as a function of p), is precisely equal to the average-sensitivity of f according to μ_p:

$$as_p(\mathsf{f}) = \frac{d\mu_p(\mathsf{f})}{dp}\,.$$

Juntas and their cores: Some functions over n binary variables as above may happen to ignore most of their input and essentially depend on only a very small, say constant, number of variables. Such functions are referred to as *juntas*. More formally, a set of variables $C \subset [n]$ is the *core* of f, if for every x,

$$\mathsf{f}(x) = \mathsf{f}(x|_C)$$

where $x|_C$ equals x on C and is otherwise 1. Furthermore, C is the (δ, p)-core of f if there exists a function f' with core C, such that,

$$\Pr_{x\sim\mu_p}\left[\mathsf{f}(x) \neq \mathsf{f}'(x)\right] \leq \delta\,.$$

A Boolean function with low total-influence is one that infrequently changes value when one of its variables is flipped at random. How can the influence be distributed among the variables? It turns out, that Boolean functions with low total-influence must have a constant-size core, namely, they are close to a junta. This is a most-insightful theorem of Friedgut [Fri98] (Theorem 3.2), which we build on herein. It states that any Boolean f has a (δ, p)-core C such that

$$|C| \leq 2^{O(as(\mathsf{f})/\delta)}\,.$$

Thus, if we allow a slight perturbation in the value of p, and since a bounded continuous function cannot have a large derivative everywhere, Russo's Lemma guarantees that a monotone Boolean function f will have low average-sensitivity. For this value of p we can apply Friedgut's Theorem, to conclude that f must be close to a junta.

One should note that this analysis in fact can serve as a proof for the following general statement: Any monotone Boolean function has a sharp threshold *unless* it is approximately determined by only a few variables. More precisely, one can prove that in any given range $[p, p+\gamma]$, a monotone Boolean function f must be close to a junta according to μ_q for some q in the range; the size of the core depending on the size of the range.

Lemma 1.2 $\forall p \in [0,1], \forall \delta, \gamma > 0, \exists q \in [p, p+\gamma]$ *s.t.* f *has a* (δ, q)*-core* C *such that* $|C| < h(p, \delta, \gamma)$.
■

1.2 Codes — Long and Biased

A *binary code* of length m is a subset

$$C \subseteq \{-1, 1\}^m$$

of strings of length m, consisting of all designated *codewords*. As mentioned above, we may view Boolean functions $\mathsf{f}\colon\{-1,1\}^n \to \{-1,1\}$ as binary vectors of dimension $m = 2^n$. Consequently, a set of Boolean functions $\mathcal{B} \subseteq \{\mathsf{f}\colon\{-1,1\}^n \to \{-1,1\}\}$ in n variables is a binary code of length $m = 2^n$.

Two parameters usually determine the quality of a binary code: (1) the *rate* of the code, $R(C) \stackrel{def}{=} \frac{1}{m}\log_2|C|$, which measures the relative entropy of C, and (2) the *distance* of the code, that is the smallest Hamming distance between two codewords. Given a set of values one wishes to encode, and a fixed distance, one would like to come up with a code whose length m is as small as possible, (i.e., the rate is as large as possible). Nevertheless, some low rate codes may enjoy other useful properties. One can apply such codes when the set of values to be encoded is very small, hence the rate is not of the utmost importance.

The *Hadamard code* is one such code, where the codewords are all characters $\{\chi_S\}_S$. Its rate is very low, with $m = 2^n$ codewords out of 2^m possible ones. Its distance is however large, being half the length, $\frac{m}{2}$.

The Long-code [BGS98] is even much sparser, containing only $n = \log m$ codewords (that is, of loglog rate). It consists of only those very particular characters $\chi_{\{i\}}$ determined by a single index i, $\chi_{\{i\}}(x) = x_i$,

$$LC = \left\{\chi_{\{i\}}\right\}_{i\in[n]}.$$

These n functions are called *dictatorship* in the influence jargon, as the value of the function is 'dictated' by a single index i.

Decoding a given string involves finding the codeword closest to it. As long as there are less than half the code's distance erroneous bit flips, unique decoding is possible since there is only one codeword within that error distance. Sometimes, the weaker notion of *list-decoding* may suffice. Here we are seeking a list of all codewords that are within a specified distance from the given string. This notion is useful when the list is guaranteed to be small. List-decoding allows a larger number of errors and helps constructing better codes, as well as plays a central role in many proofs for hardness of approximation.

Going back to the Hadamard code and the Long-code, given an arbitrary Boolean function f, the Hamming distance between f and any codeword χ_S is exactly $\frac{1-\widehat{\mathsf{f}}(S)}{2}2^n$. Since $\sum|\widehat{\mathsf{f}}(S)|^2 = 1$, there can be at most $\frac{1}{\delta^2}$ codewords that agree with f on a $\frac{1+\delta}{2}$ fraction of the points. It follows, that the Hadamard code can be list-decoded for distances up to $\frac{1-\delta}{2}2^n$. This follows through to the Long-code, being a subset of the Hadamard code.

For our purposes, however, list-decoding the Long-code is not strong enough. It is not enough that all x_i's except for those on the short list have no meaningful correlation with f. Rather, it must be the case that all of the non-listed x_i's, together, have little influence on f. In other words, f needs be close to a junta, whose variables are exactly the x_i's in the list decoding of f.

In our construction, potential codewords arise as independent sets in the *non-intersection graph* $G[n]$, defined above (Definition 1.1). Indeed, $G[n]$ has 2^n vertices, and we can think of a set of vertices of $G[n]$ as a Boolean function, by associating each vertex with an input setting in $\{-1,1\}^n$, and assigning that input -1 or $+1$ depending on whether the vertex is in or out of the set.

What are the largest independent sets in $G[n]$? One can observe that there is one for every $i \in [n]$, whose vertices correspond to all subsets S that contain i, thus containing exactly half the vertices. Viewed as a Boolean function this is just the i-th dictatorship $\chi_{\{i\}}$ which is one of the n legal codewords of the Long-code.

Other rather large independent sets exist in $G[n]$, which complicate the picture a little. Taking a few vertices out of a dictatorship independent set certainly yields an independent set. For our purposes it suffices to concentrate on maximal independent sets (ones to which no vertex can be added). Still, there are some problematic examples of large, maximal independent sets whose respective 2^n-bit string is far from all codewords: the set of all vertices S where $|S| > \frac{n}{2}$, is referred to as the *majority* independent set. Its size is very close to half the vertices, as are the dictatorships. It is easy to see, however, by a symmetry argument, that it has the same Hamming distance to all codewords (and this distance is $\approx \frac{2^n}{2}$) so there is no meaningful way of decoding it.

To solve this problem, we introduce a *bias* to the Long-code, by placing weights on the vertices of the graph $G[n]$. For every p, the weights are defined according to the p-biased product distribution:

Definition 1.2 (biased non-intersection graph) *$G_p[n]$ is a weighted graph, in which there is one vertex for each subset $S \subseteq [n]$, and where two vertices S_1 and S_2 are adjacent iff $S_1 \cap S_2 = \phi$. The weights on the vertices are as follows:*

$$\forall S \subseteq [n], \qquad \mu_p(S) = p^{|S|}(1-p)^{n-|S|} \,. \tag{1}$$

Clearly $G_{\frac{1}{2}}[n] = G[n]$ because for $p = \frac{1}{2}$ all weights are equal. Observe the manner in which we extended the notation μ_p, defined earlier as the p-biased product distribution on n-bit vectors, and now on subsets of $[n]$. The weight of each of the n dictatorship independent sets is always p. For $p < \frac{1}{2}$ and large enough n, these are the (only) largest independent sets in $G_p[n]$. In particular, the weight of the majority independent set becomes negligible.

Moreover, for $p < \frac{1}{2}$ every maximal independent set in $G_p[n]$ identifies a short list of codewords. To see that, consider a maximal independent set I in $G[n]$. The characteristic function of I —$\mathsf{f}_I(S) = -1$ if $S \in I$ and 1 otherwise— is monotone, as adding an element to a vertex S, can only decrease its neighbor set (fewer subsets S' are disjoint from it). One can apply Lemma 1.2 above to conclude that f_I must be close to a junta, for some q possibly a bit larger than p:

Corollary 1.3 *Fix $0 < p < \frac{1}{2}, \gamma > 0, \epsilon > 0$ and let I be a maximal independent set in $G_p[n]$. For some $q \in [p, p+\gamma]$, there exists $C \subset [n]$, where $|C| \leq 2^{O(1/\gamma\epsilon)}$, such that C is an (ϵ, q)-core of f_I.* ■

1.3 Extremal Set-Systems

An independent set in $G[n]$ is a family of subsets, such that every two member subsets intersect. The study of maximal intersecting families of subsets has started in the 1960s with a paper of Erdős, Ko, and Rado [EKR61]. In this classical setting, there are three parameters: $n, k, t \in \mathbb{N}$. The underlying domain is $[n]$, and one seeks the largest family of size-k subsets, every pair of which share at least t elements.

In [EKR61] it is proven that for any $k, t > 0$, and for sufficiently large n, the largest family is one that consists of all subsets that contain some t fixed elements. When n is only a constant times k this is not true. For example, the family of all subsets containing at least 3 out of 4 fixed elements is 2-intersecting, and is maximal for a certain range of values of k/n.

Frankl [Fra78], investigated the full range of values for t, k and n, and conjectured that the maximal t-intersecting family is always one of $\mathcal{A}_{i,t} \cap \binom{[n]}{k}$ where $\binom{[n]}{k}$ is the family of all size-k

subsets of $[n]$ and

$$\mathcal{A}_{i,t} \stackrel{def}{=} \{S \subseteq [n] \mid S \cap [1, \ldots, t+2i] \geq t+i\} .$$

Partial versions of this conjecture have been proven in [Fra78, FF91, Wil84]. Fortunately, the complete intersection theorem for finite sets has been settled not long ago by Ahlswede and Khachatrian [AK97].

Characterizing the largest independent sets in $G_p[n]$ amounts to studying this question for $t = 1$, yet in a *smoothed* variant. Rather than looking only at subsets of prescribed size, every subset of $[n]$ is given a weight according to μ_p, see Equation (1). Under μ_p almost all of the weight is concentrated on subsets of size roughly pn. We seek an intersecting family, largest according to this weight.

The following lemma characterizes the largest 2-intersecting families of subsets according to μ_p, in a similar manner to Alswede-Khachatrian's solution to the the Erdős-Ko-Rado question for arbitrary k.

Lemma 1.4 *Let $\mathcal{F} \subset \mathcal{P}([n])$ be 2-intersecting. For any $p < \frac{1}{2}$,*

$$\mu_p(\mathcal{F}) \leq p^{\bullet} \stackrel{def}{=} \max_i \{\mu_p(\mathcal{A}_{i,2})\} .$$

where $\mathcal{P}([n])$ denotes the power set of $[n]$. The proof is included in Appendix D.

Going back to our reduction, recall that we are transforming instances x of some NP-complete language L into graphs. Starting from a 'yes' instance ($x \in L$), the resulting graph (which is made of copies of $G_p[n]$) has an independent set whose restriction to every copy of $G_p[n]$ is a dictatorship. Hence the weight of the largest independent set in the final graph is roughly p. 'No' instances ($x \notin L$) result in a graph whose largest independent set is at most $p^{\bullet} + \epsilon$ where $p^{\bullet}$ denotes the size of the largest 2-intersecting family in $G_p[n]$. Indeed, as seen in Section 5, the final graph may contain an independent set comprising of 2-intersecting families in each copy of $G_p[n]$, regardless of whether the initial instance is a 'yes' or a 'no' instance.

Nevertheless, our analysis shows that any independent set in $G_p[n]$ whose size is even marginally larger than the largest 2-intersecting family of subsets, identifies an index $i \in [n]$. This 'assignment' of value i per copy of $G_p[n]$ can then serve to prove that the starting instance x is a 'yes' instance.

In summary, the source of our inapproximability factor comes from the gap between sizes of maximal 2-intersecting and 1-intersecting families. This factor is $\frac{1-p^{\bullet}}{1-p}$, being the ratio between the sizes of the vertex covers that are the complements of the independent sets discussed above. The value of p is constrained by additional technical complications stemming from the structure imposed by the PCP theorem.

1.4 Stronger PCP theorems and hardness of approximation

The PCP theorem was originally stated and proven in the context of probabilistic checking of proofs. However, it has a clean interpretation as a constraint satisfaction problem (sometimes referred to as Label-Cover), which we now formulate explicitly. There are two sets of non-Boolean variables, X and Y. The variables take values in finite domains R_x and R_y respectively. For some of the pairs (x, y), $x \in X$ and $y \in Y$, there is a constraint $\pi_{x,y}$. A constraint specifies which values for x and y will satisfy it. Furthermore, all constraints must have the 'projection' property. Namely, for every x-value there is only one possible y-value that together would satisfy the constraint. An enhanced version of the PCP theorem states that

Theorem 1.5 (The PCP Theorem [AS98, ALM$^+$98, Raz98]) *Given as input a system of constraints $\{\pi_{x,y}\}$ as above, it is NP-hard to decide whether*

- *There is an assignment to X, Y that satisfies all of the constraints.*
- *There is no assignment that satisfies more than an $|R_x|^{-\Omega(1)}$ fraction of the constraints.*

A general scheme for proving hardness of approximation was developed in [BGS98, Hås01, Hås99]. The equivalent of this scheme in our setting, would be to construct a copy of the intersection graph for every variable in $X \cup Y$. The copies would then be further connected according to the constraints between the variables, in a straightforward way.

It turns out that such a construction can only work if the constraints between the x, y pairs in the PCP theorem are *extremely* restricted. The important 'bijection-like' parameter is as follows: given any value for one of the variables, how many values for the other variable will still satisfy the constraint. In projection constraints, a value for the x variable has only one possible extension to a value for the y variable; but a value for the y variable may leave many possible values for x. In contrast, a significant part of our construction is devoted to getting symmetric two-variable constraints where values for one variable leave 1 or 2 possibilities for the second variable, *and vice versa*. It is the precise structure of these constraints that limits p to being at most $\frac{3-\sqrt{5}}{2}$.

In fact, our construction proceeds by transformations on graphs rather than on constraint satisfaction systems. We employ a well-known reduction [FGL$^+$96] converting the constraint satisfaction system of Theorem 1.5 to a graph made of cliques that are further connected. We refer to such a graph as co-partite because it is the complement of a multi-partite graph. The reduction asserts that in this graph it is NP-hard to approximate the maximum independent set, with some additional technical requirements. The major step is to transform this graph into a new co-partite graph that has a crucial additional property, as follows. Every two cliques are either totally disconnected, or, they induce a graph such that the co-degree of every vertex is either 1 or 2. This is analogous to the 'bijection-like' parameter of the constraints discussed above.

1.5 Minimum Vertex Cover

Let us now briefly describe the history of the Minimum Vertex Cover problem. There is a simple greedy algorithm that approximates Minimum Vertex Cover to within a factor of 2 as follows: Greedily obtain a maximal matching in the graph, and let the vertex cover consist of both vertices at the ends of each edge in the matching. The resulting vertex-set covers all the edges and is no more than twice the size of the smallest vertex cover. Using the best currently known algorithmic tools doesn't help much in this case, and the best known algorithm gives an approximation factor of $2 - o(1)$ [Hal02, BYE85, MS83].

As to hardness results, the previously best known hardness result was due to Håstad [Hås01] who showed that it is NP-hard to approximate Minimum Vertex Cover to within a factor of $\frac{7}{6}$. Let us remark that both Håstad's result and the result presented herein hold for graphs of bounded degree. This follows simply because the graph resulting from our reduction is of bounded degree.

1.6 Organization of the Paper

The reduction is described in Section 2. In Section 2.1 we define a specific variant of the gap independent set problem called hIS and show it NP-hard. This encapsulates all one needs to know – for the purpose of our proof – of the PCP theorem. Section 2.2 describes the reduction from an instance of hIS to Minimum Vertex Cover. The reduction starts out from a graph G and constructs from it the final graph $G_{\mathcal{B}}^{\mathbb{C}}$. The section ends with the (easy) proof of completeness of the reduction. Namely, that if $\mathrm{IS}(G) = m$ then $G_{\mathcal{B}}^{\mathbb{C}}$ contains an independent set whose relative size is roughly $p \approx 0.38$.

The main part of the proof is the proof of soundness. Namely, proving that if the graph G is a 'no' instance, then the largest independent set in $G_{\mathcal{B}}^{\mathbb{C}}$ has relative size at most $< p^{\bullet} + \varepsilon \approx 0.159$. Section 3 surveys the necessary technical background; and Section 4 contains the proof itself. Finally, Section 5 contains some examples showing that the analysis of our construction is tight.

2 The Construction

In this section we describe our construction. We first define a specific gap variant of the Maximum Independent Set problem. The NP-hardness of this problem follows directly from known results, and it encapsulates all one needs to know about PCP for our proof. We then describe the reduction from this problem to Minimum Vertex Cover.

2.1 Co-partite Graphs and h-Clique-Independence

Consider the following type of graph,

Definition 2.1 *An* (m, r)-co-partite graph $G = \langle M \times R, E\rangle$ *is a graph constructed of $m = |M|$ cliques each of size $r = |R|$, hence the edge set of G satisfies*

$$\forall i \in M, j_1 \neq j_2 \in R, \quad (\langle i, j_1\rangle, \langle i, j_2\rangle) \in E\,.$$

Such a graph is the complement of an m-partite graph, whose parts have r vertices each. It follows from the proof of [FGL+96], that it is NP-hard to approximate the Maximum Independent Set specifically on (m, r)-co-partite graphs.

Next, consider the following strengthening of the concept of an independent set,

Definition 2.2 *For any graph $G = (V, E)$, define*

$$\mathrm{IS}_h(G) \stackrel{def}{=} \max\{\, |I| \;\mid\; I \subseteq V \text{ contains no clique of size } h\}\,.$$

The gap-h-Clique-Independent-Set Problem *(or $h\mathrm{IS}(r, \epsilon, h)$ for short) is as follows:*

Instance: *An (m, r)-co-partite graph G.*

Problem: *Distinguish between the following two cases:*

- $\mathrm{IS}(G) = m$.
- $\mathrm{IS}_h(G) \leq \epsilon m$.

Note that for $h = 2$, $\mathrm{IS}_2(G) = \mathrm{IS}(G)$, and this becomes the usual gap-Independent-Set problem. Nevertheless, by a standard reduction, one can show that this problem is still hard, as long as r is large enough compared to h:

Theorem 2.1 *For any $h, \epsilon > 0$, the problem $h\mathrm{IS}(r, \epsilon, h)$ is NP-hard, as long as $r \geq (\frac{h}{\epsilon})^c$ for some constant c.*

A complete derivation of this theorem from the PCP theorem can be found in Appendix B.

2.2 The Reduction

In this section we present our reduction from $h\mathrm{IS}(r, \varepsilon_0, h)$ to Minimum Vertex Cover by constructing, from any given (m, r)-co-partite graph G, a graph $G_{\mathcal{B}}^{\mathbb{C}}$. Our main theorem is as follows,

Theorem 2.2 *For any $\varepsilon > 0$, and $p < p_{\max} = \frac{3-\sqrt{5}}{2}$, for large enough h, l_{T} and small enough ε_0 (see Definition 2.3 below): Given an (m, r)-co-partite graph $G = (M \times R, E)$, one can construct, in polynomial time, a graph $G_{\mathcal{B}}^{\mathbb{C}}$ so that:*

$$\begin{aligned} \mathrm{IS}(G) = m \quad &\Longrightarrow \quad \mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) \geq p - \varepsilon \\ \mathrm{IS}_h(G) < \varepsilon_0 \cdot m \quad &\Longrightarrow \quad \mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) < p^\bullet + \varepsilon \qquad \textit{where } p^\bullet = \max(p^2,\, 4p^3 - 3p^4)\,. \end{aligned}$$

As an immediate corollary we obtain,

Corollary 2.3 (Independent-Set) *Let $p < p_{\max} = \frac{3-\sqrt{5}}{2}$. For any constant $\varepsilon > 0$, given a weighted graph G, it is NP-hard to distinguish between:*

Yes: $\quad \mathrm{IS}(G) > p - \varepsilon$

No: $\quad \mathrm{IS}(G) < p^\bullet + \varepsilon$

In case $p \leq \frac{1}{3}$, $p^\bullet$ reads p^2 and the above asserts that it is NP-hard to distinguish between $\mathcal{I}(G_{\mathcal{B}}^{\mathbb{C}}) \approx p = \frac{1}{3}$ and $\mathcal{I}(G_{\mathcal{B}}^{\mathbb{C}}) \approx p^2 = \frac{1}{9}$ and the gap between the sizes of the minimum vertex cover in the 'yes' and 'no' cases approaches $\frac{1-p^2}{1-p} = 1 + p$, yielding a hardness-of-approximation factor of $\frac{4}{3}$ for Minimum Vertex Cover. Our main result follows immediately,

Theorem 1.1 *Given a graph G, it is NP-hard to approximate Minimum Vertex Cover to within any factor smaller than $10\sqrt{5} - 21 \approx 1.3606$.*

Proof: For $\frac{1}{3} < p < p_{\max}$, direct computation shows that $p^\bullet = 4p^3 - 3p^4$, thus it is NP-hard to distinguish between the case $G_{\mathcal{B}}^{\mathbb{C}}$ has a vertex cover of size $1 - p + \epsilon$ and the case $G_{\mathcal{B}}^{\mathbb{C}}$ has a vertex cover of size at least $1 - 4p^3 + 3p^4 - \epsilon$ for any $\epsilon > 0$. Minimum Vertex Cover is thus shown hard to approximate to within a factor approaching

$$\frac{1 - 4(p_{\max})^3 + 3(p_{\max})^4}{1 - p_{\max}} = 1 + p_{\max} + (p_{\max})^2 - 3(p_{\max})^3 = 10\sqrt{5} - 21 \approx 1.36068...$$

■

Before we turn to the proof of the main theorem, let us introduce some parameters needed during the course of the proof. It is worthwhile to note here that the particular values chosen for these parameters are insignificant. They are merely chosen so as to satisfy some assertions through the course of the proof, nevertheless, most importantly, they are all independent of $r = |R|$. Once the proof has demonstrated that assuming a $(p^\bullet + \varepsilon)$-weight independent set in

$G_{\mathcal{B}}^{\mathbb{C}}$, there must be a set of weight ε_0 in G that contains no h-clique, one can set r to be large enough so as to imply NP-hardness of $h\mathrm{IS}(r, \varepsilon_0, h)$, which thereby implies NP-hardness for the appropriate gap-Independent-Set problem. This argument is valid due to the fact that none of the parameters of the proof is related to r.

Definition 2.3 (Parameter Setting) *Given $\varepsilon > 0$ and $p < p_{\max}$, let us set the following parameters:*

- *Let $0 < \gamma < p_{\max} - p$ be such that, $(p+\gamma)^{\bullet} - p^{\bullet} < \frac{1}{4}\varepsilon$.*
- *Choosing h: We choose h to accommodate applications of Friedgut's theorem (Theorem 3.2 below), a Sunflower Lemma and a pigeon-hole principle. Let $\Gamma(p, \delta, k)$ be the function defined in Theorem 3.2, and let $\Gamma_*(k, d)$ be the function defined in the Sunflower Lemma (Theorem 4.8 below). Set*

$$h_0 = \sup_{q \in [p, p_{\max}]} \left(\Gamma(q, \tfrac{1}{16}\varepsilon, \tfrac{2}{\gamma}) \right)$$

  and let $\eta = \frac{1}{16 h_0} \cdot p^{5h_0}$, $h_1 = \lceil \frac{2}{\gamma \cdot \eta} \rceil + h_0$, $h_s = 1 + 2^{2h_0} \cdot \sum_{k=0}^{h_0} \binom{h_1}{k}$, and $h = \Gamma_*(h_1, h_s)$.
- *Fix $\varepsilon_0 = \frac{1}{32} \cdot \varepsilon$.*
- *Fix $l_{\mathsf{T}} = \max(4 \ln \frac{2}{\varepsilon},\ (h_1)^2)$.*

Remarks. The value of γ is well defined because the function taking p to $p^{\bullet} = \max(p^2, 4p^3 - 3p^4)$ is a continuous function of p. The supremum $\sup_{q \in [p, p_{\max}]} \left(\Gamma(q, \frac{1}{16}\varepsilon, \frac{2}{\gamma}) \right)$ in the definition of h_0 is bounded, because $\Gamma(q, \frac{1}{16}\varepsilon, \frac{2}{\gamma})$ is a continuous function of q, see Theorem 3.2. Both r and l_{T} remain fixed while the size of the instance $|G|$ increases to infinity, so without loss of generality we can assume that $l_{\mathsf{T}} \cdot r \ll m$.

Constructing the Final Graph $G_{\mathcal{B}}^{\mathbb{C}}$

Let us denote the set of vertices of G by $V = M \times R$.

The constructed graph $G_{\mathcal{B}}^{\mathbb{C}}$ will depend on a parameter $l \stackrel{def}{=} 2 l_{\mathsf{T}} \cdot r$.

Consider the family $\mathcal{B}$ of all sets of size l of V:

$$\mathcal{B} = \binom{V}{l} = \{ B \subset V \mid |B| = l \} .$$

Let us refer to each such $B \in \mathcal{B}$ as a *block*. The intersection of an independent set $\mathcal{I}_G \subset V$ in G with any $B \in \mathcal{B}$, $\mathcal{I}_G \cap B$, can take 2^l distinct forms, namely all subsets of B. If $|\mathcal{I}_G| = m$ then expectedly $|\mathcal{I}_G \cap B| = l \cdot \frac{m}{mr} = 2 l_{\mathsf{T}}$ hence for almost all B it is the case that $|\mathcal{I}_G \cap B| > l_{\mathsf{T}}$. Let us consider for each block B its *block-assignments*,

$$R_B \stackrel{def}{=} \left\{ \mathsf{a}\colon B \to \{\mathsf{T}, \mathsf{F}\} \;\middle|\; |\mathsf{a}^{-1}(\mathsf{T})| \geq l_{\mathsf{T}} \right\} .$$

Every block-assignment $\mathsf{a} \in R_B$ supposedly corresponds to some independent set $\mathcal{I}_G$, and assigns T to exactly all vertices of B that are in $\mathcal{I}_G$, that is, where $\mathsf{a}^{-1}(\mathsf{T}) = \mathcal{I}_G \cap B$. Two block-assignments are adjacent in $G_{\mathcal{B}}$ if they are surely not referring to the same independent set. In this case they will be said to be *inconsistent*. Thus $\mathsf{a} \neq \mathsf{a}' \in R_B$ are inconsistent.

Consider a pair of blocks B_1, B_2 that intersect on $\hat{B} = B_1 \cap B_2$ with $|\hat{B}| = l - 1$. For a block-assignment $\mathsf{a}_1 \in R_{B_1}$, let us denote by $\mathsf{a}_1|_{\hat{B}} \colon \hat{B} \to \{\mathsf{T}, \mathsf{F}\}$ the restriction of a_1 to $\hat{B}$, namely, where $\forall v \in \hat{B},\ \mathsf{a}_1|_{\hat{B}}(v) = \mathsf{a}_1(v)$. Block assignments $\mathsf{a}_1 \in R_{B_1}$ and $\mathsf{a}_2 \in R_{B_2}$ are possibly referring to the same independent set only if $\mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}}$. If also $B_1 = \hat{B} \cup \{v_1\}$ and $B_2 = \hat{B} \cup \{v_2\}$ such that v_1, v_2 are adjacent in G, $\mathsf{a}_1, \mathsf{a}_2$ are consistent only if they don't both assign T to v_1, v_2 respectively. In summary, every block-assignment $\mathsf{a}_1 \in R_{B_1}$ is consistent with (and will not be adjacent to) at most *two* block-assignments in R_{B_2}.

Let us formally construct the graph $G_{\mathcal{B}} = (V_{\mathcal{B}}, E_{\mathcal{B}})$,

Definition 2.4 *Define the graph $G_{\mathcal{B}} = (V_{\mathcal{B}}, E_{\mathcal{B}})$, with vertices for all block-assignments to every block $B \in \mathcal{B}$,*

$$V_{\mathcal{B}} = \bigcup_{B \in \mathcal{B}} R_B$$

and edges for every pair of block-assignments that are clearly inconsistent,

$$E_{\mathcal{B}} = \bigcup_{\langle v_1, v_2 \rangle \in E,\ \hat{B} \in \binom{V}{l-1}} \left\{ \langle \mathsf{a}_1, \mathsf{a}_2 \rangle \in R_{\hat{B} \cup \{v_1\}} \times R_{\hat{B} \cup \{v_2\}} \,\middle|\, \mathsf{a}_1|_{\hat{B}} \neq \mathsf{a}_2|_{\hat{B}} \quad or \quad \mathsf{a}_1(v_1) = \mathsf{a}_2(v_2) = \mathsf{T} \right\}$$
$$\bigcup_{B} \{ \langle \mathsf{a}_1, \mathsf{a}_2 \rangle \mid \mathsf{a}_1, \mathsf{a}_2 \in R_B \} .$$

Note that $|R_B|$ is the same for all $B \in \mathcal{B}$, so for $r' = |R_B|$ and $m' = |\mathcal{B}|$, the graph $G_{\mathcal{B}}$ is (m', r')-co-partite.

The (almost perfect) completeness of the reduction from G to $G_{\mathcal{B}}$, can be easily proven:

Proposition 2.4 $\mathrm{IS}(G) = m \quad \Longrightarrow \quad \mathrm{IS}(G_{\mathcal{B}}) \geq m' \cdot (1 - \varepsilon)$.

Proof: Let $\mathcal{I}_G \subset V$ be an independent set in G, $|\mathcal{I}| = m = \frac{1}{r}|V|$. Let $\mathcal{B}'$ consist of all l-sets $B \in \mathcal{B} = \binom{V}{l}$ that intersect $\mathcal{I}_G$ on at least l_{T} elements $|B \cap \mathcal{I}_G| \geq l_{\mathsf{T}}$. The probability that this does not happen is (see Proposition E.1) $\Pr_{B \in \mathcal{B}}[B \notin \mathcal{B}'] \leq 2e^{-\frac{2 l_{\mathsf{T}}}{8}} \leq \varepsilon$. For a block $B \in \mathcal{B}'$, let $\mathsf{a}_B \in R_B$ be the characteristic function of $\mathcal{I}_G \cap B$:

$$\forall v \in B \qquad \mathsf{a}_B(v) \stackrel{def}{=} \begin{cases} \mathsf{T} & v \in \mathcal{I}_G \\ \mathsf{F} & v \notin \mathcal{I}_G \end{cases}$$

The set $\mathcal{I} = \{ \mathsf{a}_B \mid B \in \mathcal{B}' \}$ is an independent set in $G_{\mathcal{B}}$, of size $m' \cdot (1 - \varepsilon)$. ■

The Final Graph

We now define our final graph $G_{\mathcal{B}}^{\mathbb{C}}$, consisting of the same blocks as $G_{\mathcal{B}}$, but where each block is not a clique but rather a copy of the non-intersection graph $G_p[n]$, for $n = |R_B|$, as defined in the introduction (Definition 1.2).

Vertices and Weights: $G_{\mathcal{B}}^{\mathbb{C}} = \left\langle V_{\mathcal{B}}^{\mathbb{C}}, E_{\mathcal{B}}^{\mathbb{C}}, \Lambda \right\rangle$ has a block of vertices $V_{\mathcal{B}}^{\mathbb{C}}[B]$ for every $B \in \mathcal{B}$, where vertices in each block B correspond to the non-intersection graph $G_p[n]$, for $n = |R_B|$. We identify every vertex of $V_{\mathcal{B}}^{\mathbb{C}}[B]$ with a subset of R_B, that is,

$$V_{\mathcal{B}}^{\mathbb{C}}[B] = \mathcal{P}(R_B) .$$

$V_{\mathcal{B}}^{\mathbb{C}}$ consists of one such block of vertices for each $B \in \mathcal{B}$,

$$V_{\mathcal{B}}^{\mathbb{C}} = \bigcup_{B \in \mathcal{B}} V_{\mathcal{B}}^{\mathbb{C}}[B] .$$

Note that we take the block-assignments to be distinct, hence, subsets of them are distinct, and $V_{\mathcal{B}}^{\mathbb{C}}$ is a disjoint union of $V_{\mathcal{B}}^{\mathbb{C}}[B]$ over all $B \in \mathcal{B}$.

Let Λ_B, for each block $B \in \mathcal{B}$, be the distribution over the vertices of $V_{\mathcal{B}}^{\mathbb{C}}[B]$, as defined in Definition 1.2. Namely, we assign each vertex F a probability according to μ_p:

$$\Lambda_B(F) = \mu_p^{R_B}(F) = p^{|F|}(1-p)^{|R_B \setminus F|} .$$

Finally, the probability distribution Λ assigns equal probability to every block: For any $F \in V_{\mathcal{B}}^{\mathbb{C}}[B]$

$$\Lambda(F) \stackrel{def}{=} |\mathcal{B}|^{-1} \cdot \Lambda_B(F)$$

Edges. We have edges between every pair of $F_1 \in V_{\mathcal{B}}^{\mathbb{C}}[B_1]$ and $F_2 \in V_{\mathcal{B}}^{\mathbb{C}}[B_2]$ if in the graph $G_{\mathcal{B}}$ there is a complete bipartite graph between these sets, i.e.

$$E_{\mathcal{B}}^{\mathbb{C}} = \left\{ \langle F_1, F_2 \rangle \in V_{\mathcal{B}}^{\mathbb{C}}[B_1] \times V_{\mathcal{B}}^{\mathbb{C}}[B_2] \;\middle|\; E_{\mathcal{B}} \supseteq F_1 \times F_2 \right\}$$

In particular, there are edges within a block, i.e. when $B_1 = B_2$, iff $F_1 \cap F_2 = \phi$ (formally, this follows from the definition because the vertices of R_B form a clique in $G_{\mathcal{B}}$, and $G_{\mathcal{B}}$ has no self edges).

This completes the construction of the graph $G_{\mathcal{B}}^{\mathbb{C}}$. We have,

Proposition 2.5 *For any fixed $p, l > 0$, the graph $G_{\mathcal{B}}^{\mathbb{C}}$ is polynomial-time constructible given input G.* ■

A simple-to-prove, nevertheless crucial, property of $G_{\mathcal{B}}^{\mathbb{C}}$ is that every independent set[1] can be monotonely extended,

Proposition 2.6 *Let $\mathcal{I}$ be an independent set of $G_{\mathcal{B}}^{\mathbb{C}}$: If $F \in \mathcal{I} \cap V_{\mathcal{B}}^{\mathbb{C}}[B]$, and $F \subset F' \in V_{\mathcal{B}}^{\mathbb{C}}[B]$, then $\mathcal{I} \cup \{F'\}$ is also an independent set.* ■

We conclude this section by proving completeness of the reduction:

Lemma 2.7 (Completeness) $\mathrm{IS}(G) = m \quad \Longrightarrow \quad \mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) \geq p - \varepsilon$.

Proof: By proposition 2.4, if $\mathrm{IS}(G) = m$ then $\mathrm{IS}(G_{\mathcal{B}}) \geq m'(1-\varepsilon)$. In other words, there is an independent set $\mathcal{I}_{\mathcal{B}} \subset V_{\mathcal{B}}$ of $G_{\mathcal{B}}$ whose size is $|\mathcal{I}_{\mathcal{B}}| \geq m' \cdot (1-\varepsilon)$. Let $\mathcal{I}_0 = \{ \{\mathsf{a}\} \mid \mathsf{a} \in \mathcal{I}_{\mathcal{B}} \}$ be the independent set consisting of all singletons of $\mathcal{I}_{\mathcal{B}}$, and let $\mathcal{I}$ be $\mathcal{I}_0$'s monotone closure. The set $\mathcal{I}$ is also an independent set due to Proposition 2.6 above. It remains to observe that the weight within each block of the family of all sets containing a fixed $\mathsf{a} \in \mathcal{I}_{\mathcal{B}}$, is p. ■

[1] An independent set in the intersection graph never contains the empty-set vertex, because it has a self loop.

3 Technical Background

In this section we describe our technical tools, formally defining and stating theorems that were already described in the introduction. As described in the introduction, these theorems come from distinct fields, in particular harmonic analysis of Boolean functions and extremal set theory.

For the rest of the paper, we will adopt the notation of extremal set theory as follows. A family of subsets of a finite set R will usually be denoted by $\mathcal{F} \subseteq \mathcal{P}(R)$, and member subsets by $F, H \in \mathcal{F}$. We represent a Boolean function $\mathsf{f} : \{-1,1\}^n \to \{-1,1\}$, according to its alternative view as a family of subsets

$$\mathcal{F} = \{ F \in \mathcal{P}(R) \,|\, \mathsf{f}(\sigma_F) = -1 \},$$

where σ_F is the vector with -1 on coordinates in F, and 1 otherwise.

3.1 A Family's Core

A family of subsets $\mathcal{F} \subset \mathcal{P}(R)$ is said to be a *junta* with *core* $C \subset R$, if a subset $F \in \mathcal{P}(R)$ is determined to be in or out of $\mathcal{F}$ only according to its intersection with C (no matter whether other elements are in or out of F). Formally, C is the core of $\mathcal{F}$ if,

$$\{ F \in \mathcal{P}(R) \,|\, F \cap C \in \mathcal{F} \} = \mathcal{F}.$$

A given family $\mathcal{F}$, does not necessarily have a small core C. However, there might be another family $\mathcal{F}'$ with core C, and that approximates $\mathcal{F}$ quite accurately, up to some δ:

Definition 3.1 (Core) *A set $C \subseteq R$ is said to be a (δ, p)-core of the family $\mathcal{F} \subseteq \mathcal{P}(R)$, if there exists a junta $\mathcal{F}' \subseteq \mathcal{P}(R)$ with core C such that $\mu_p(\mathcal{F} \triangle \mathcal{F}') < \delta$.*

The family $\mathcal{F}'$ that best approximates $\mathcal{F}$ on its core, consists of the subsets $F \in \mathcal{P}(C)$ whose extension to R intersects more than half of $\mathcal{F}$,

$$[\mathcal{F}]_C^{\frac{1}{2}} \stackrel{def}{=} \left\{ F \in \mathcal{P}(C) \;\middle|\; \Pr_{F' \in \mu_p^{R \setminus C}} \left[F \cup F' \in \mathcal{F} \right] > \frac{1}{2} \right\}.$$

Consider the *Core-Family*, defined as the family of all subsets $F \in \mathcal{P}(C)$, for which $\frac{3}{4}$ of their extension to R, i.e. $\frac{3}{4}$ of $\{ F' \,|\, F' \cap C = F \}$, reside in $\mathcal{F}$:

Definition 3.2 (Core-Family) *For a set of elements $C \subset R$, define,*

$$[\mathcal{F}]_C^{\frac{3}{4}} \stackrel{def}{=} \left\{ F \in \mathcal{P}(C) \;\middle|\; \Pr_{F' \in \mu_p^{R \setminus C}} \left[F \cup F' \in \mathcal{F} \right] > \frac{3}{4} \right\}.$$

By simple averaging, it turns out that if C is a (δ, p)-core for $\mathcal{F}$, this family approximates $\mathcal{F}$ almost as well as the best family C.

Lemma 3.1 *If C is a (δ, p)-core of $\mathcal{F}$, then $\mu_p^C \left([\mathcal{F}]_C^{\frac{3}{4}} \right) \geq \mu_p^R(\mathcal{F}) - 4\delta$.*

Proof: Clearly, $[\mathcal{F}]_C^{\frac{1}{2}} \supseteq [\mathcal{F}]_C^{\frac{3}{4}}$. Let

$$\mathcal{F}_{\frac{1}{2}} = \left\{ F \;\middle|\; F \cap C \in [\mathcal{F}]_C^{\frac{1}{2}} \right\} \qquad \mathcal{F}_{\frac{3}{4}} = \left\{ F \;\middle|\; F \cap C \in [\mathcal{F}]_C^{\frac{3}{4}} \right\},$$

and let $\mathcal{F}' = \mathcal{F}_{\frac{1}{2}} \setminus \mathcal{F}_{\frac{3}{4}}$. We will show

$$\mu((\mathcal{F} \triangle \mathcal{F}_{\frac{3}{4}}) \cap \mathcal{F}') \leq 3\mu((\mathcal{F} \triangle \mathcal{F}_{\frac{1}{2}}) \cap \mathcal{F}') \tag{2}$$

and thus

$$\begin{aligned} \mu(\mathcal{F} \triangle \mathcal{F}_{\frac{3}{4}}) \leq \mu(((\mathcal{F} \triangle \mathcal{F}_{\frac{3}{4}}) \cap \mathcal{F}') \cup ((\mathcal{F} \triangle \mathcal{F}_{\frac{3}{4}}) \cap \overline{\mathcal{F}'})) &\leq 3\mu((\mathcal{F} \triangle \mathcal{F}_{\frac{1}{2}}) \cap \mathcal{F}') + \mu((\mathcal{F} \triangle \mathcal{F}_{\frac{3}{4}}) \cap \overline{\mathcal{F}'}) \\ &= 3\mu((\mathcal{F} \triangle \mathcal{F}_{\frac{1}{2}}) \cap \mathcal{F}') + \mu((\mathcal{F} \triangle \mathcal{F}_{\frac{1}{2}}) \cap \overline{\mathcal{F}'}) \leq 4\delta \end{aligned}$$

where the first line follows from (2) and second line holds because $\mathcal{F}_{\frac{1}{2}} = \mathcal{F}_{\frac{3}{4}}$ outside $\mathcal{F}'$.

To prove (2), fix $F \in [\mathcal{F}]_C^{\frac{1}{2}} \setminus [\mathcal{F}]_C^{\frac{3}{4}}$, and denote

$$\rho = \Pr_{F' \in \mu_p^{R \setminus C}} \left[F \cup F' \in \mathcal{F} \right].$$

Clearly $\frac{1}{2} < \rho \leq \frac{3}{4}$ so $(1 - \rho) \geq \rho/3$. For every $F' \subseteq R \setminus C$, the subset $F \cup F'$ is always in $\mathcal{F}_{\frac{1}{2}}$ and not in $\mathcal{F}_{\frac{3}{4}}$, so

$$\Pr_{F' \in \mu_p^{R \setminus C}} \left[F \cup F' \in \mathcal{F} \triangle \mathcal{F}_{\frac{1}{2}} \right] \;=\; 1 - \rho \;\geq\; \frac{\rho}{3} = \frac{1}{3} \cdot \Pr_{F' \in \mu_p^{R \setminus C}} \left[F \cup F' \in \mathcal{F} \triangle \mathcal{F}_{\frac{3}{4}} \right].$$

■

Influence and Sensitivity

Let us now define influence and average sensitivity for families of subsets. Assume a family of subsets $\mathcal{F} \subseteq \mathcal{P}(R)$. The *influence* of an element $e \in R$,

$$\mathsf{influence}_e^p(\mathcal{F}) \stackrel{def}{=} \Pr_{F \in \mu_p} [\text{ exactly one of } F \cup \{e\},\ F \setminus \{e\} \text{ is in } \mathcal{F}].$$

The *total-influence* or *average sensitivity* of $\mathcal{F}$ with respect to μ_p, denoted $\mathsf{as}_p(\mathcal{F})$, is the sum of the influences of all elements in R,

$$\mathsf{as}_p(\mathcal{F}) \stackrel{def}{=} \sum_{e \in R} \mathsf{influence}_e^p(\mathcal{F}).$$

Friedgut's theorem states that if the average sensitivity of a family is small, then it has a small (δ, p)-core:

Theorem 3.2 (Theorem 4.1 in [Fri98]) *Let $0 < p < 1$ be some bias, and $\delta > 0$ be any approximation parameter. Consider any family $\mathcal{F} \subset \mathcal{P}(R)$, and let $k = \mathsf{as}_p(\mathcal{F})$. There exists a function $\Gamma(p, \delta, k) \leq (c_p)^{k/\delta}$, where c_p is a constant depending only on p, such that $\mathcal{F}$ has a (δ, p)-core C, with $|C| \leq \Gamma(p, \delta, k)$.* ■

Remark. We rely on the fact that the constant c_p above is bounded by a continuous function of p. The dependence of c_p on p follows from Friedgut's p-biased equivalent of the Bonami-Beckner inequality. In particular, there is a parameter $1 < \tau < 2$ whose precise value depends on p as follows: it must satisfy $(\tau - 1)p^{2/\tau - 1} > 1 - 3\tau/4$. Clearly τ is a continuous (bounded) function of p.

A family of subsets $\mathcal{F} \subseteq \mathcal{P}(R)$ is *monotone* if for every $F \in \mathcal{F}$, for all $F' \supset F$, $F' \in \mathcal{F}$. We will use the following easy fact,

Proposition 3.3 *For a monotone family $\mathcal{F} \subseteq \mathcal{P}(R)$, $\mu_p(\mathcal{F})$ is a monotone non-decreasing function of p.*

For a simple proof of this proposition, see Appendix C.

Interestingly, for monotone families, the rate at which μ_p increases with p, is exactly equal to the average-sensitivity:

Theorem 3.4 (Russo-Margulis Identity [Mar74, Rus82]) *Let $\mathcal{F} \subseteq \mathcal{P}(R)$ be a monotone family. Then,*

$$\frac{d\mu_p(\mathcal{F})}{dp} = \mathsf{as}_p(\mathcal{F})$$

$\blacksquare$

For a simple proof of this identity, see Appendix C.

3.2 Maximal Intersecting Families

Recall from the introduction that a monotone family distinguishes a small core of elements, that almost determine it completely. Next, we will show that a monotone family that is of large enough weight, and is also *intersecting*, must exhibit one *distinguished* element in its core. This element will consequently serve for establishing consistency between distinct families.

Definition 3.3 *A family $\mathcal{F} \subset \mathcal{P}(R)$ is* t-intersecting, *for $t \geq 1$, if*

$$\forall F_1, F_2 \in \mathcal{F}, \quad |F_1 \cap F_2| \geq t\,.$$

For $t = 1$ such a family is referred to simply as intersecting.

Let us first consider the following natural generalization for a pair of families,

Definition 3.4 (Cross-Intersecting) *Two families $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{P}(R)$ are* cross-intersecting *if for every $F_1 \in \mathcal{F}_1$ and $F_2 \in \mathcal{F}_2$, $F_1 \cap F_2 \neq \phi$.*

Two families cannot be too large and still remain cross-intersecting,

Proposition 3.5 *Let $p \leq \frac{1}{2}$, and let $\mathcal{F}_1, \mathcal{F}_2 \subseteq \mathcal{P}(R)$ be two families of subsets for which $\mu_p(\mathcal{F}_1) + \mu_p(\mathcal{F}_2) > 1$. Then $\mathcal{F}_1, \mathcal{F}_2$ are not cross-intersecting.*

Proof: We can assume that $\mathcal{F}_1, \mathcal{F}_2$ are monotone, as their monotone closures must also be cross-intersecting. Since μ_p, for a monotone family, is non-decreasing with respect to p (see Proposition 3.3), it is enough to prove the claim for $p = \frac{1}{2}$.

For a given subset F denote its complement by $F^{\mathsf{c}} = R \setminus F$. If there was some $F \in \mathcal{F}_1 \cap \mathcal{F}_2$ for which $F^{\mathsf{c}} \in \mathcal{F}_1$ or $F^{\mathsf{c}} \in \mathcal{F}_2$, then clearly the families would not be cross-intersecting. Yet if such a subset $F \in \mathcal{F}_1 \cap \mathcal{F}_2$ does not exist, then the sum of sizes of $\mathcal{F}_1, \mathcal{F}_2$ would be bounded by 1. ■

It is now easy to prove that if $\mathcal{F}$ is monotone and intersecting, then the same holds for the core-family $[\mathcal{F}]_C^{\frac{3}{4}}$ that is (see Definition 3.2) the threshold approximation of $\mathcal{F}$ on its core C,

Proposition 3.6 *Let $\mathcal{F} \subseteq \mathcal{P}(R)$, and let $C \subseteq R$.*

- *If $\mathcal{F}$ is monotone then $[\mathcal{F}]_C^{\frac{3}{4}}$ is monotone.*
- *If $\mathcal{F}$ is intersecting, and $p \leq \frac{1}{2}$, then $[\mathcal{F}]_C^{\frac{3}{4}}$ is intersecting.*

Proof: The first assertion is immediate. For the second assertion, assume by way of contradiction, a pair of non-intersecting subsets $F_1, F_2 \in [\mathcal{F}]_C^{\frac{3}{4}}$ and observe that the families

$$\{F \in \mathcal{P}(R \setminus C) \mid F \cup F_1 \in \mathcal{F}_1\} \quad \text{and} \quad \{F \in \mathcal{P}(R \setminus C) \mid F \cup F_2 \in \mathcal{F}_2\}$$

both have weight $> \frac{3}{4}$, and by Proposition 3.5, cannot be cross intersecting. ■

An intersecting family whose weight is larger than that of a maximal 2- intersecting family, must contain two subsets that intersect on a unique element $e \in R$.

Definition 3.5 (Distinguished Element) *For a monotone and intersecting family $\mathcal{F} \subseteq \mathcal{P}(R)$, an element $e \in R$ is said to be* distinguished *if there exist $F^\flat, F^\sharp \in \mathcal{F}$ such that*

$$F^\flat \cap F^\sharp = \{e\}.$$

The distinguished element itself is not unique, a fact that is irrelevant to our analysis as we choose an arbitrary one. Clearly, an intersecting family has a distinguished element if and only if it is not 2-intersecting. We next establish a weight criterion for an intersecting family to have a distinguished element. Recall that $p_{\max} = \frac{3-\sqrt{5}}{2}$. For each $p < p_{\max}$, define $p^\bullet$ to be

Definition 3.6

$$\forall p < p_{\max}, \qquad p^\bullet \stackrel{def}{=} \max(p^2,\, 4p^3 - 3p^4).$$

This maps each p to the size of the maximal 2-intersecting family, according to μ_p. For a proof of such a bound we venture into the field of extremal set theory, where maximal intersecting families have been studied for some time. This study was initiated by Erdős, Ko, and Rado [EKR61], and has seen various extensions and generalizations. The corollary above is a generalization to μ_p of what is known as the Complete Intersection Theorem for finite sets, that was proven in [AK97]. Frankl [Fra78] defined the following families:

$$\mathcal{A}_{i,t} \stackrel{def}{=} \{F \in \mathcal{P}([n]) \mid F \cap [1, t+2i] \geq t+i\}$$

which are easily seen to be t-intersecting for $0 \leq i \leq \frac{n-t}{2}$; and conjectured the following theorem that was finally proven by Ahlswede and Khachatrian [AK97]:

Theorem 3.7 ([AK97]) *Let $\mathcal{F} \subseteq \binom{[n]}{k}$ be t-intersecting. Then,*

$$|\mathcal{F}| \leq \max_{0 \leq i \leq \frac{n-t}{2}} \left| \mathcal{A}_{i,t} \cap \binom{[n]}{k} \right|$$

■

Our analysis requires the extension of this statement to families of subsets that are not restricted to a specific size k, and where $t = 2$. Let us denote $\mathcal{A}_i \stackrel{def}{=} \mathcal{A}_{i,2}$. The following lemma (mentioned in the introduction) follows from the above theorem, and will be proven in Appendix D.

Lemma 1.4 *Let $\mathcal{F} \subset \mathcal{P}([n])$ be 2-intersecting. For any $p < \frac{1}{2}$,*

$$\mu_p(\mathcal{F}) \leq \max_i \{\mu_p(\mathcal{A}_i)\}.$$

Furthermore, when $p \leq \frac{1}{3}$, this maximum is attained by $\mu_p(\mathcal{A}_0) = p^2$, and for $\frac{1}{3} < p < p_{\max}$ by $\mu_p(\mathcal{A}_1) = 4p^3 - 3p^4$. Having defined $p^\bullet = \max(p^2, 4p^3 - 3p^4)$ for every $p < p_{\max}$, we thus have

Corollary 3.8 *If $\mathcal{F} \subset \mathcal{P}(R)$ is 2-intersecting, then $\mu_p(\mathcal{F}) \leq p^\bullet$, provided $p < p_{\max}$.*

The proof of this corollary can also be found in Appendix D.

4 Soundness

This section is the heart, and most technical part, of the proof of correctness, proving the construction is *sound*, that is, that if $G_{\mathcal{B}}^{\mathbb{C}}$ has a large independent set, then G has a large h-clique–free set.

Lemma 4.1 (Soundness) $\mathrm{IS}(G_{\mathcal{B}}^{\mathbb{C}}) \geq p^\bullet + \varepsilon \quad \Longrightarrow \quad \mathrm{IS}_h(G) \geq \varepsilon_0 \cdot m.$

Proof Sketch: Assuming an independent set $\mathcal{I} \subset V_{\mathcal{B}}^{\mathbb{C}}$ of weight $\Lambda(\mathcal{I}) \geq p^\bullet + \varepsilon$, we consider for each block $B \in \mathcal{B}$ the family $\mathcal{I}[B] = \mathcal{I} \cap V_{\mathcal{B}}^{\mathbb{C}}[B]$.

The first step (Lemma 4.2) is to find, for a non-negligible fraction of the blocks $\mathcal{B}_q \subseteq \mathcal{B}$, a small core of permissible block-assignments, and in it, one distinguished block-assignment to be used later to form a large h-clique–free set in G. This is done by showing that for every $B \in \mathcal{B}_q$, $\mathcal{I}[B]$ has both significant weight and low average-sensitivity. This, not necessarily true for p, is asserted for some slightly shifted value $q \in (p, p + \gamma)$. Utilizing Friedgut's theorem, we deduce the existence of a small core for $\mathcal{I}[B]$. Then, utilizing an Erdős-Ko-Rado-type bound on the maximal size of a 2-intersecting family, we find a distinguished block-assignment for each $B \in \mathcal{B}_q$.

The next step is to focus on one (e.g. random) $l - 1$ sub-block $\hat{B} \in \binom{V}{l-1}$, and consider its extensions $\hat{B} \cup \{v\}$ for $v \in V = M \times R$, that represent the initial graph G. The distinguished block-assignments of those blocks that are in $\mathcal{B}_q$ will serve to identify a large set in V.

The final most delicate part of the proof, is Lemma 4.6, asserting that the distinguished block-assignments of the blocks extending $\hat{B}$ must identify an h-clique–free set as long as $\mathcal{I}$ is an independent set. Indeed, since they all share the same $(l-1)$-sub-block $\hat{B}$, the edge constraints these blocks impose on one another will suffice to conclude the proof.

After this informal sketch, let us now turn to the formal proof of Lemma 4.1.

Proof: Let then $\mathcal{I} \subset V_{\mathcal{B}}^{\mathbb{C}}$ be an independent set of size $\Lambda(\mathcal{I}) \geq p^\bullet + \varepsilon$, and denote, for each $B \in \mathcal{B}$,

$$\mathcal{I}[B] \stackrel{def}{=} \mathcal{I} \cap V_{\mathcal{B}}^{\mathbb{C}}[B].$$

The fractional size of $\mathcal{I}[B]$ within $V_{\mathcal{B}}^{\mathbb{C}}[B]$, according to Λ_B, is $\Lambda_B(\mathcal{I}[B]) = \mu_p(\mathcal{I}[B])$.

Assume w.l.o.g. that $\mathcal{I}$ is maximal,

Observation. $\mathcal{I}[B]$, for any $B \in \mathcal{B}$, is monotone and intersecting.

Proof: It is intersecting, as $G_{\mathcal{B}}^{\mathbb{C}}$ has edges connecting vertices corresponding to non-intersecting subsets, and it is monotone due to maximality (see Proposition 2.6). ∎

The first step in our proof is to find, for a significant fraction of the blocks, a small core, and in it one distinguished block-assignment. Recall from Definition 3.5, that an element $\mathsf{a} \in C$ would be distinguished for a family $[\mathcal{I}[B]]_C^{\frac{3}{4}} \subseteq \mathcal{P}(C)$ if there are two subsets $F^\flat, F^\sharp \in [\mathcal{I}[B]]_C^{\frac{3}{4}}$ whose intersection is exactly $F^\flat \cap F^\sharp = \{\mathsf{a}\}$.

Theorem 3.2 implies that a family has a small core only if the family has low average-sensitivity, which is not necessarily the case here. To overcome this, let us use an extension of Corollary 1.3, which would allow us to assume some q slightly larger than p, for which a large fraction of the blocks have a low average-sensitivity, and thus a small core. By the weight of the family being large, it follows that there must be a distinguished block-assignment in that core.

Lemma 4.2 *There exists some $q \in [p, p_{\max})$, and a set of blocks $\mathcal{B}_q \subseteq \mathcal{B}$ whose size is $|\mathcal{B}_q| \geq \frac{1}{4}\varepsilon \cdot |\mathcal{B}|$, such that for all $B \in \mathcal{B}_q$:*

1. *$\mathcal{I}[B]$ has an $(\frac{1}{16}\varepsilon, q)$-core, $\mathsf{Core}[B] \subset R_B$, of size $|\mathsf{Core}[B]| \leq h_0$.*
2. *The core-family $[\mathcal{I}[B]]_{\mathsf{Core}[B]}^{\frac{3}{4}}$ has a distinguished element $\dot{\mathsf{a}}[B] \in \mathsf{Core}[B]$.*

Proof: We will find a value $q \in [p, p_{\max})$ and a set of blocks $\mathcal{B}_q \subseteq \mathcal{B}$ such that for every $B \in \mathcal{B}_q$, $\mathcal{I}[B]$ is of large weight and low average sensitivity, according to μ_q. We will then proceed to show that this implies the above properties. First consider blocks whose intersection with $\mathcal{I}$ has weight not much lower than the expectation,

$$\mathcal{B}' \stackrel{def}{=} \left\{ B \in \mathcal{B} \;\middle|\; \Lambda_B(\mathcal{I}[B]) > p^\bullet + \frac{1}{2}\varepsilon \right\}.$$

By a simple averaging argument, it follows that $|\mathcal{B}'| \geq \frac{1}{2}\varepsilon \cdot |\mathcal{B}|$, as otherwise

$$\Lambda(\mathcal{I}) \cdot |\mathcal{B}| = \sum_{B \in \mathcal{B}} \Lambda_B(\mathcal{I}[B]) \leq \frac{1}{2}\varepsilon\,|\mathcal{B}| + \sum_{B \notin \mathcal{B}'} \Lambda_B(\mathcal{I}[B]) < \frac{1}{2}\varepsilon\,|\mathcal{B}| + \sum_{B \notin \mathcal{B}'} (p^\bullet + \frac{1}{2}\varepsilon) \leq (p^\bullet + \varepsilon) \cdot |\mathcal{B}|$$

Since μ_p is non-decreasing with p (see Proposition 3.3), and since the value of $\gamma < p_{\max} - p$ was chosen so that for every $q \in [p, p+\gamma]$, $p^\bullet + \frac{1}{4}\varepsilon > q^\bullet$, we have for every block $B \in \mathcal{B}'$,

$$\mu_q(\mathcal{I}[B]) \geq \mu_p(\mathcal{I}[B]) > p^\bullet + \frac{1}{2}\varepsilon > q^\bullet + \frac{1}{4}\varepsilon\,. \tag{3}$$

The family $\mathcal{I}[B]$, being monotone, cannot have high average sensitivity according to μ_q for many values of q, so by allowing an increase of at most γ, the set

$$\mathcal{B}_q \stackrel{def}{=} \left\{ B \in \mathcal{B}' \;\middle|\; \mathsf{as}_q(\mathcal{I}[B]) \leq \frac{2}{\gamma} \right\}$$

must be large for some $q \in [p, p+\gamma]$:

Proposition 4.3 *There exists $q \in [p, p+\gamma]$ so that $|\mathcal{B}_q| \geq \frac{1}{4}\varepsilon \cdot |\mathcal{B}|$.*

Proof: Consider the average, within $\mathcal{B}'$, of the size of $\mathcal{I}[B]$ according to μ_q

$$\mu_q[\mathcal{B}'] \stackrel{def}{=} \left|\mathcal{B}'\right|^{-1} \cdot \sum_{B\in\mathcal{B}'} \mu_q(\mathcal{I}[B])$$

and apply a version of Lagrange's Mean-Value Theorem: The derivative of $\mu_q[\mathcal{B}']$ as a function of q is

$$\frac{d\mu_q[\mathcal{B}']}{dq} = \left|\mathcal{B}'\right|^{-1} \cdot \sum_{B\in\mathcal{B}'} \frac{d\mu_q}{dq}(\mathcal{I}[B]) = \left|\mathcal{B}'\right|^{-1} \cdot \sum_{B\in\mathcal{B}'} \mathsf{as}_q(\mathcal{I}[B])$$

where the last equality follows from the Russo-Margulis identity (Lemma 3.4). Therefore, there must be some $q \in [p, p+\gamma]$ for which $\frac{d\mu_q[\mathcal{B}']}{dq} \leq \frac{1}{\gamma}$, as otherwise $\mu_{p+\gamma}[\mathcal{B}'] > 1$ which is impossible. It follows that at least half of the blocks in $\mathcal{B}'$ have $\mathsf{as}_q(\mathcal{I}[B]) \leq \frac{2}{\gamma}$. We have $|\mathcal{B}_q| \geq \frac{1}{2}|\mathcal{B}'| \geq \frac{1}{4}\varepsilon\,|\mathcal{B}|$. ■

Fix then $q \in [p, p+\gamma]$, to be as in the proposition above, so that $|\mathcal{B}_q| \geq \frac{1}{4}\varepsilon \cdot |\mathcal{B}|$. We next show that the properties claimed by the lemma, indeed hold for all blocks in $\mathcal{B}_q$. The first property, namely that $\mathcal{I}[B]$ has an $(\frac{1}{16}\varepsilon, q)$-core, denoted $\mathsf{Core}[B] \subset R_B$, of size $|\mathsf{Core}[B]| \leq h_0$, is immediate from Theorem 3.2, plugging in the average sensitivity of $\mathcal{I}[B]$, and by definition of $h_0 = \sup_{q\in[p,p_{\max}]} \Gamma(q, \frac{1}{16}\,\varepsilon, \frac{2}{\gamma})$, see Definition 2.3.

Having found a core for $\mathcal{I}[B]$, consider the core-family approximating $\mathcal{I}[B]$ on $\mathsf{Core}[B]$, (see Definition 3.2), denoted by

$$\mathcal{CF}_B \stackrel{def}{=} [\mathcal{I}[B]]^{\frac{3}{4}}_{\mathsf{Core}[B]} = \left\{ F \in \mathcal{P}\left(\mathsf{Core}[B]\right) \;\middle|\; \Pr_{F'\in\mu_p^{R\setminus\mathsf{Core}[B]}}\left[F \cup F' \in \mathcal{I}[B]\right] > \frac{3}{4} \right\}.$$

By Proposition 3.6, since $\mathcal{I}[B]$ is monotone and intersecting, so is $\mathcal{CF}_B$. Moreover, Corollary 3.1 asserts that

$$\mu_q(\mathcal{CF}_B) > \mu_q(\mathcal{I}[B]) - 4 \cdot \frac{\varepsilon}{16} > q^\bullet$$

where the second inequality follows from inequality (3), stating that $\mu_q(I[B]) > q^\bullet + \frac{1}{4}\varepsilon$ for any $B \in \mathcal{B}_q$. We can now utilize the bound on the maximal size of a 2-intersecting family (see Corollary 3.8), to deduce that $\mathcal{CF}_B$ is too large to be 2-intersecting, and must distinguish an element $\dot{\mathsf{a}} \in \mathsf{Core}[B]$, i.e. contain two subsets $F^\sharp, F^\flat \in \mathcal{CF}_B$ that intersect on exactly that block-assignment, $F^\sharp \cap F^\flat = \{\dot{\mathsf{a}}\}$. This completes the proof of Lemma 4.2. ■

Let us now fix q as guaranteed by Lemma 4.2 above. The following implicit definitions appeared in the above proof, and will be used later as well,

Definition 4.1 (Core, Core-Family, Distinguished Block-Assignment) *Let $B \in \mathcal{B}_q$.*

- *B's* core*, denoted $\mathsf{Core}[B] \subset R_B$, is an arbitrary smallest $(\frac{1}{16}\varepsilon, q)$-core of $\mathcal{I}[B]$.*
- *B's* core-family*, is the core-family on B's core (see Definition 3.2), denoted $\mathcal{CF}_B = [\mathcal{I}[B]]^{\frac{3}{4}}_{\mathsf{Core}[B]}$.*
- *B's* distinguished block-assignment*, is an arbitrary distinguished element of $\mathcal{CF}_B$, denoted $\dot{\mathsf{a}}[B] \in \mathsf{Core}[B]$.*

Let us further define for each block $B \in \mathcal{B}_q$, the set of all block-assignments of B that have influence larger than $\eta = \frac{1}{16h_0} \cdot p^{8h_0}$:

Definition 4.2 (Extended Core) *For $B \in \mathcal{B}$, let the* extended core *of B be*

$$\mathit{ECore}[B] \stackrel{def}{=} \mathit{Core}[B] \cup \{\mathsf{a} \in R_B \mid \mathsf{influence}^q_{\mathsf{a}}(\mathcal{I}[B]) \geq \eta\} \ .$$

The extended core is not much larger than the core, because the total sum of influences of elements in R_B, is bounded for every $B \in \mathcal{B}_q$, by $\mathsf{as}_q(\mathcal{I}[B]) \leq \frac{2}{\gamma}$,

$$|\mathsf{ECore}[B]| \leq h_0 + \frac{\mathsf{as}_q(\mathcal{I}[B])}{\eta} \leq h_0 + \lceil \tfrac{2}{\gamma \cdot \eta} \rceil = h_1 \ .$$

Consider now an $(l-1)$-sub-block $\hat{B} \in \binom{V}{l-1}$. The set of l-blocks that extend $\hat{B}$ can be thought of as a copy of G. The next step in our proof is to identify one such sub-block, and a set of blocks extending it (say $\hat{B} \cup \{v_1\}, \ldots, \hat{B} \cup \{v_m\}$) so that the corresponding subset of vertices $\{v_1, \ldots, v_m\} = V_{\hat{B}} \subset V$ is h-clique–free. Members of $V_{\hat{B}}$ are determined in a delicate way as follows. For each block $\hat{B} \cup \{v\} \in \mathcal{B}_q$, if the distinguished block-assignment of that block assigns T to v, then v is put in $V_{\hat{B}}$ ($V_{\hat{B}}$ is formally defined in Definition 4.4). We show in Proposition 4.5 that an appropriate random selection of $\hat{B}$ implies that $V_{\hat{B}}$ is sufficiently large. Then, in Lemma 4.6 we analyze the cores and distinguished block-assignments of the blocks $\hat{B} \cup \{v_1\}, \ldots, \hat{B} \cup \{v_m\}$, and deduce that the set $V_{\hat{B}}$ must be h-clique free.

In Figure 1 the top two lines represent the block assignments of B_1 and the two bottom lines represent the block assignments of B_2. The lines are labeled by T and F to indicate the value assigned to v_1 (resp. v_2) by block-assignments on that line. The center line represents the sub-block assignments $R_{\hat{B}}$. The block assignments are aligned so that all 5 in the same column agree on the assignr

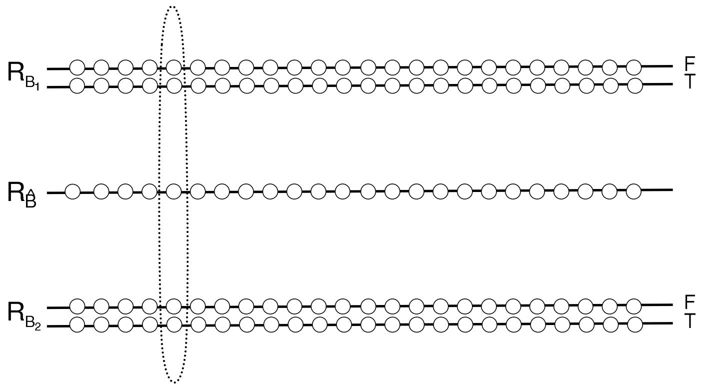

Figure 1: Aligned pairs of block assignments

The key is that only block-assignments that are in the same column can be consistent, thus a pair of block-assignments $\mathsf{a}_1 \in R_{B_1}$ and $\mathsf{a}_2 \in R_{B_2}$ are consistent only if their restriction to $\hat{B}$ is equal (i.e. they are in the same column), and assuming v_1, v_2 are adjacent in G they must not both assign T to v_1 and v_2 respectively.

We must attend a small technical issue before we continue. It would be undesired to have both block-assignments in a given pair influential in $\mathcal{I}[B]$ for this would mean that the structure of $\mathcal{I}[B]$ is not preserved when reduced to $\hat{B}$. Thus, besides requiring that many of the blocks $\hat{B} \cup \{v\}$ extending $\hat{B}$ reside in $\mathcal{B}_q$ (and have a well-defined core and distinguished block-assignment), we need them to be *preserved* by $\hat{B}$:

Definition 4.3 (Preservation) *Let $B \in \mathcal{B}$, and let $\hat{B} \subset B$, $|\hat{B}| = l-1$. Let us denote by $\mathsf{a}|_{\hat{B}}$ the restriction to $\hat{B}$ of a block-assignment $\mathsf{a} \in R_B$. We say that $\hat{B}$* preserves *B, if there is no pair of block-assignments $\mathsf{a}_1 \neq \mathsf{a}_2 \in \mathit{ECore}[B]$ with $\mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}}$.*

It is almost always the case that $\hat{B}$ preserves $\hat{B} \cup \{v\}$:

Proposition 4.4

$$\forall B \in \mathcal{B} \quad |\{v \in B \mid B \setminus \{v\} \textit{ does not preserve } B\}| < \frac{(h_1)^2}{2}.$$

Proof: Each pair of block-assignments $\mathsf{a}_1, \mathsf{a}_2 \in \mathsf{ECore}[B]$ can cause at most one $\hat{B}$ to not preserve B, and for any block $B \in \mathcal{B}_q$, $|\mathsf{ECore}[B]| \leq h_1$; consequently, the number of $\hat{B}$ not preserving B is at most $\binom{h_1}{2} < \frac{(h_1)^2}{2}$. ■

The last step before identifying the required $\hat{B}$ is to see that a distinguished block-assignment for a block $\hat{B} \cup \{v\}$ is useful for constructing an h-clique–free subset in G, if it assigns T to v. Hence, for each $\hat{B}$ we consider the following set $V_{\hat{B}} \subset V$:

Definition 4.4 *Let $V_{\hat{B}} \subseteq V$ be:*

$$V_{\hat{B}} \stackrel{def}{=} \left\{ v \in V \setminus \hat{B} \;\middle|\; B = \hat{B} \cup \{v\} \in \mathcal{B}_q \quad \textit{and} \quad \hat{B} \textit{ preserves } B \quad \textit{and} \quad \dot{\mathsf{a}}[B](v) = \mathsf{T} \right\}.$$

It follows from the definition of $V_{\hat{B}}$, that if $v_1, v_2 \in V_{\hat{B}}$ are connected by an edge in G, then the distinguished block-assignments of $B_1 = \hat{B} \cup \{v_1\}$ and $B_2 = \hat{B} \cup \{v_2\}$ are connected by an edge in the graph $G_{\mathcal{B}}$, $\langle \dot{\mathsf{a}}[B_1], \dot{\mathsf{a}}[B_2] \rangle \in E_{\mathcal{B}}$ (see Definition 2.4). Next, let us identify a sub-block $\hat{B}$, for which $V_{\hat{B}}$ is large:

Proposition 4.5 *There exists $\hat{B} \in \binom{V}{l-1}$, with $\left|V_{\hat{B}}\right| \geq \varepsilon_0 \cdot m$.*

Proof: Observe that

$$\Pr_{\hat{B},\, v \in V \setminus \hat{B}} \left[v \in V_{\hat{B}} \right] \geq \frac{1}{4}\varepsilon \cdot \Pr_{B,\, v \in B} \left[v \in V_{B \setminus \{v\}} \;\middle|\; B \in \mathcal{B}_q \right] \geq \frac{1}{4}\varepsilon \cdot \frac{1}{4r}$$

where the first inequality follows from Proposition 4.3 asserting $\mathcal{B}_q \geq \frac{1}{4}\varepsilon\,|\mathcal{B}|$. The second inequality is a consequence of the fact that for any $\mathsf{a} \in R_B$, there are at least $l_{\mathsf{T}} = \frac{l}{2r}$ elements $v \in B$ with $\mathsf{a}(v) = \mathsf{T}$; and at most $\frac{(h_1)^2}{2}$ $(l-1)$-blocks $\hat{B} \subset B$ not preserving B; hence, conditioned on $B \in \mathcal{B}_q$, the probability of $v \in V_{\hat{B}}$ is at least $\frac{1}{2r} - \frac{(h_1)^2}{2l} \geq \frac{1}{4r}$ as $l \geq 2(h_1)^2 \cdot r$.

This inequality shows that there is at least one $\hat{B}$ for which $\Pr_{v \in V \setminus \hat{B}} \left[v \in V_{\hat{B}} \right] \geq \frac{\varepsilon}{16r}$, hence, $\left|V_{\hat{B}}\right| \geq \frac{1}{16r}\varepsilon \cdot \left|V \setminus \hat{B}\right| \geq \frac{1}{32}\varepsilon \cdot m$, as $\frac{|V \setminus \hat{B}|}{r} > \frac{1}{2}m$, because $|\hat{B}| = l-1 \ll \frac{1}{2}|V|$, see Definition 2.3. ■

Finally, we establish $\mathrm{IS}_h(G) \geq \varepsilon_0 \cdot m$ by proving,

Lemma 4.6 *The set $V_{\hat{B}}$ contains no clique of size h.*

Proof: (of Lemma 4.6) Assume, by way of contradiction, that there exists a clique over vertices $v_1, \ldots, v_h \in V_{\hat{B}}$. We show that, for $B_i = \hat{B} \cup \{v_i\}$, the set $\cup_{i \in [h]} \mathcal{I}[B_i]$ is not an independent set. In fact, we explicitly find two of these blocks, B_{i_1}, B_{i_2}, such that $\mathcal{I}[B_{i_1}] \cup \mathcal{I}[B_{i_2}]$ is not an independent set.

Analyzing consistency between blocks $\hat{B} \cup \{v_i\}$ leads us to consider the common sub-block $\hat{B}$, and the sub-block-assignments that are restrictions of block-assignments in R_{B_i} to $\hat{B}$. The $(l-1)$-*block-assignments* of $\hat{B} \in \binom{V}{l-1}$, are defined to be

$$R_{\hat{B}} \stackrel{def}{=} \left\{ \mathsf{a} \colon \hat{B} \to \{\mathsf{T}, \mathsf{F}\} \right\}.$$

A block-assignment $\mathsf{a} \in R_{B_i}$ has a natural restriction to $\hat{B}$, denoted $\mathsf{a}|_{\hat{B}} \in R_{\hat{B}}$, where $\forall v \in \hat{B}$, $\mathsf{a}|_{\hat{B}}(v) = \mathsf{a}(v)$.

For the remaining analysis, let us name the three important entities regarding each block B_i, for $i \in [h]$: B_i's distinguished block-assignment, the core of B_i, and the extended core of B_i,

$$\dot{\mathsf{a}}_i \stackrel{def}{=} \dot{\mathsf{a}}[B_i] \qquad C_i \stackrel{def}{=} \mathsf{Core}[B_i] \qquad E_i \stackrel{def}{=} \mathsf{ECore}[B_i]$$

and their natural restrictions to $\hat{B}$ (where the natural restriction of a set is the set comprising the restrictions of its elements),

$$\hat{\mathsf{a}}_i \stackrel{def}{=} \dot{\mathsf{a}}_i|_{\hat{B}} \qquad \hat{C}_i \stackrel{def}{=} C_i|_{\hat{B}} \qquad \hat{E}_i \stackrel{def}{=} E_i|_{\hat{B}}.$$

Now, recall the core-family $\mathcal{CF}_{B_i}$, which is the family of subsets, over the core of each B_i, whose extension in $\mathcal{I}[B_i]$ has weight at least $\frac{3}{4}$. For each block B_i, $i \in [h]$, $\dot{\mathsf{a}}_i$ being distinguished implies a pair of subsets

$$F_i^\flat, F_i^\sharp \in \mathcal{CF}_{B_i} \text{ so that } F_i^\flat \cap F_i^\sharp = \{\dot{\mathsf{a}}_i\}$$

Let their natural restriction to $\hat{B}$ be

$$\hat{F}_i^\flat \stackrel{def}{=} F_i^\flat|_{\hat{B}} \qquad \hat{F}_i^\sharp \stackrel{def}{=} F_i^\sharp|_{\hat{B}}$$

and note that, as $\hat{B}$ preserves every B_i, it follows that, for all $i \in [h]$,

$$\hat{F}_i^\flat \cap \hat{F}_i^\sharp = \{\hat{\mathsf{a}}_i\}. \tag{4}$$

Our first goal is to identify two blocks B_{i_1} and B_{i_2} whose core-families look the same in the following sense,

Proposition 4.7 *There exist $i_1 \neq i_2 \in [h]$, such that, denoting $\mathbf{\Delta} = \hat{E}_{i_1} \cap \hat{E}_{i_2}$,*

1. $\hat{C}_{i_1} \cap \mathbf{\Delta} = \hat{C}_{i_2} \cap \mathbf{\Delta}$
2. $\hat{F}_{i_1}^\flat \cap \mathbf{\Delta} = \hat{F}_{i_2}^\flat \cap \mathbf{\Delta}$
3. $\hat{F}_{i_1}^\sharp \cap \mathbf{\Delta} = \hat{F}_{i_2}^\sharp \cap \mathbf{\Delta}$

Proof: Our proof begins by applying the following Sunflower-Lemma over the sets $\hat{E}_i$:

Theorem 4.8 ([ER60]) *There exists some integer function $\Gamma_*(k,d)$ (not depending on $|R|$), such that for any $\mathcal{F} \subset \binom{R}{k}$, if $|\mathcal{F}| \geq \Gamma_*(k,d)$, there are d distinct sets $F_1, \ldots, F_d \in \mathcal{F}$, such that, let $\mathbf{\Delta} \stackrel{def}{=} F_1 \cap \ldots \cap F_d$, the sets $F_i \setminus \mathbf{\Delta}$ are pairwise disjoint.* ∎

The sets $F_1, .., F_d$ are called a Sunflower, or a $\mathbf{\Delta}$-system. This statement can easily be extended to families in which each subset is of size *at most* k.

We apply this lemma for $R = R_{\hat{B}}$, and $\mathcal{F} = \{\hat{E}_1, .., \hat{E}_h\}$. Recall (definition 2.3) we have fixed $h > \Gamma_*(h_1, h_s)$, hence Theorem 4.8 implies there exists some $J \subseteq [h]$, $|J| = h_s$, such that

$$\left\{\hat{E}_i \setminus \mathbf{\Delta}\right\}_{i \in J} \text{ are pairwise disjoint for } \mathbf{\Delta} \stackrel{def}{=} \bigcap_{i \in J} \hat{E}_i\,.$$

Consider, for each $i \in J$, the triplet $\left\langle \hat{C}_i \cap \mathbf{\Delta}, \quad \hat{F}_i^\flat \cap \mathbf{\Delta}, \quad \hat{F}_i^\sharp \cap \mathbf{\Delta} \right\rangle$, and note that, since $\hat{F}_i^\flat, \hat{F}_i^\sharp \subseteq \hat{C}_i$ the number of possible triplets is at most

$$\begin{aligned} \left|\left\{ \left\langle \hat{C} \cap \mathbf{\Delta}, \hat{F}^\flat \cap \mathbf{\Delta}, \hat{F}^\sharp \cap \mathbf{\Delta} \right\rangle \;\middle|\; |\hat{C}| \le h_0,\, \hat{F}^\flat, \hat{F}^\sharp \subseteq \hat{C} \right\}\right| &\le \sum_{k=0}^{h_0} \binom{h_1}{k} \cdot 2^{h_0} \cdot 2^{h_0} \\ &< h_s = |J| \end{aligned}$$

(recall we have set (definition 2.3) $h_s = 1 + 2^{2h_0} \cdot \sum_{k=0}^{h_0} \binom{h_1}{k}$). Therefore, by the pigeon-hole principle, there must be some $i_1, i_2 \in J$ for which

$$\left\langle \hat{C}_{i_1} \cap \mathbf{\Delta},\ \hat{F}_{i_1}^\flat \cap \mathbf{\Delta},\ \hat{F}_{i_1}^\sharp \cap \mathbf{\Delta} \right\rangle = \left\langle \hat{C}_{i_2} \cap \mathbf{\Delta},\ \hat{F}_{i_2}^\flat \cap \mathbf{\Delta},\ \hat{F}_{i_2}^\sharp \cap \mathbf{\Delta} \right\rangle$$

■

¿From now on let us assume w.l.o.g. that $i_1 = 1$, $i_2 = 2$, and continue to denote $\mathbf{\Delta} = \hat{E}_1 \cap \hat{E}_2$. We will arrive at a contradiction by finding an edge between the blocks B_1, B_2, specifically, by finding two extensions, one of $F_1^\flat$ in $\mathcal{I}[B_1]$, and another of $F_2^\sharp$ in $\mathcal{I}[B_2]$, all of whose block-assign

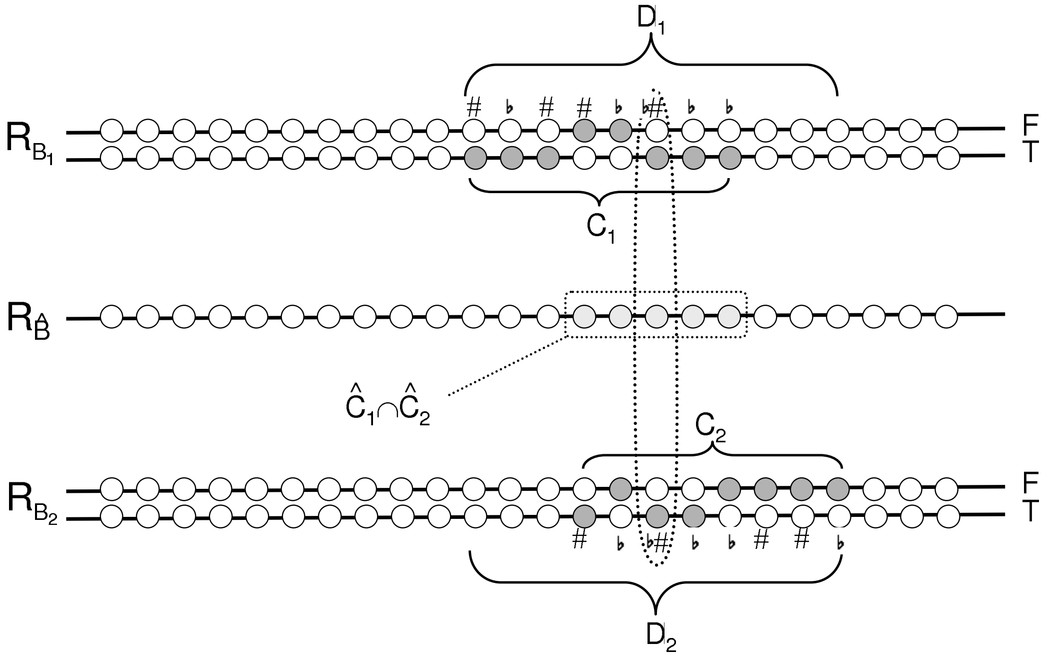

Figure 2: Cores and Distinguished Block-Assignments

Figure 2 can be helpful in keeping track of the important entities in the rest of the proof. Recall that two block assignments are consistent *only* if they are in the same column and are not both in the T row. The darker circles represent members of the core (C_1 or C_2) and note that there is at most one darker circle in each T/F pair (due to preservation). The block-assignments in $F_1^\flat$ and $F_1^\sharp$ are labeled $\flat$ and $\sharp$. The distinguished block-assignments are labeled by both $\flat$ and $\sharp$, and they assign T to v_1, v_2 respectively. The dashed rectangle borders the intersection of $\hat{C}_1$ with $\hat{C}_2$, which is a subset of $\mathbf{\Delta}$ and is where the restrictions of $F_1^\sharp, F_1^\flat$ are equal to those of $F_2^\sharp, F_2^\flat$.

As a first step, let us prove that the block-assignments in $F_1^\flat$ and $F_2^\sharp$ are pairwise inconsistent:

Proposition 4.9 $\left\langle F_1^\flat, F_2^\sharp \right\rangle \in E_{\mathcal{B}}^{\mathbb{L}}$.

Proof: We need to prove that for all $\mathsf{a}_1 \in F_1^\flat, \mathsf{a}_2 \in F_2^\sharp$, $\langle \mathsf{a}_1, \mathsf{a}_2 \rangle \in E_{\mathcal{B}}$. If $\langle \mathsf{a}_1, \mathsf{a}_2 \rangle \notin E_{\mathcal{B}}$, it must be that $\mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}} \in \hat{F}_1^\flat \cap \hat{F}_2^\sharp \subseteq \hat{E}_1 \cap \hat{E}_2 = \mathbf{\Delta}$. B_1 and B_2 are chosen in Proposition 4.7 so that $\hat{F}_1^\flat \cap \mathbf{\Delta} = \hat{F}_2^\flat \cap \mathbf{\Delta}$ and $\hat{F}_1^\sharp \cap \mathbf{\Delta} = \hat{F}_2^\sharp \cap \mathbf{\Delta}$. Consequently $\mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}} \in \hat{F}_1^\flat \cap \hat{F}_1^\sharp \cap \mathbf{\Delta} = \hat{F}_2^\flat \cap \hat{F}_2^\sharp \cap \mathbf{\Delta}$, however (4) asserts that the only block-assignment in these two intersections is the distinguished one, hence $\hat{\mathsf{a}}_1 = \mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}} = \hat{\mathsf{a}}_2$. Since $\hat{B}$ preserves both B_1 and B_2, $\mathsf{a}_1 = \dot{\mathsf{a}}_1$ and $\mathsf{a}_2 = \dot{\mathsf{a}}_2$. However, $\langle \dot{\mathsf{a}}_1, \dot{\mathsf{a}}_2 \rangle \in E_{\mathcal{B}}$ (recall Definition 2.4), as both $\dot{\mathsf{a}}_1, \dot{\mathsf{a}}_2$ assign T to v_1, v_2 respectively and $\langle v_1, v_2 \rangle \in E$. ∎

It may well be that $F_1^\flat \notin \mathcal{I}[B_1]$ and $F_2^\sharp \notin \mathcal{I}[B_2]$, thus the proposition above is only a first step towards a contradiction. Nevertheless, we know that $F_1^\flat \in \mathcal{CF}_{B_1} = [\mathcal{I}[B_1]]_{\mathsf{Core}[B_1]}^{\frac{3}{4}}$ means that $\frac{3}{4}$ of $\left\{ F \in \mathcal{P}\left(R_{B_1}\right) \mid F \cap \mathsf{Core}[B_1] = F_1^\flat \right\}$ are in $\mathcal{I}[B_1]$; and likewise for $F_2^\sharp$. In what follows, we utilize this large volume of $\frac{3}{4}$ to find extensions of these sets, that are in $\mathcal{I}$, yet are connected by an edge in $E_{\mathcal{B}}^{\mathbb{L}}$.

Let us partition the set of $(l-1)$-block assignments of $R_{\hat{B}}$ into the important ones, which are restrictions of block-assignments in the cores of B_1 or B_2, and the rest,

$$\hat{D} = \hat{C}_1 \cup \hat{C}_2 \quad \text{and} \quad \hat{R} = R_{\hat{B}} \setminus \hat{D}$$

which immediately partitions the block-assignments of R_{B_1} and R_{B_2}, according to whether their restriction falls within $\hat{D}$:

$$D_1 = \left\{ \mathsf{a} \in R_{B_1} \;\middle|\; \mathsf{a}|_{\hat{B}} \in \hat{D} \right\} \quad \text{and} \quad R_1 = R_{B_1} \setminus D_1$$

and similarly for R_{B_2},

$$D_2 = \left\{ \mathsf{a} \in R_{B_2} \;\middle|\; \mathsf{a}|_{\hat{B}} \in \hat{D} \right\} \quad \text{and} \quad R_2 = R_{B_2} \setminus D_2$$

Proposition 4.10 $|D_1| \leq 4h_0$ *and* $|D_2| \leq 4h_0$.

Proof: Simply note that $|D_1|, |D_2| \leq 2|\hat{D}| \leq 2(|\hat{C}_1| + |\hat{C}_2|) \leq 2(|C_1| + |C_2|) = 4h_0$. ∎

Notice that $F_1^\flat \in \mathcal{P}\left(C_1\right) \subseteq \mathcal{P}\left(D_1\right)$ and $F_2^\sharp \in \mathcal{P}\left(C_2\right) \subseteq \mathcal{P}\left(D_2\right)$, hence it suffices to exhibit two subsets $H_1 \in \mathcal{P}\left(R_1\right)$ and $H_2 \in \mathcal{P}\left(R_2\right)$ all of whose block-assignments are pairwise-inconsistent, and so that $F_1^\flat \cup H_1 \in \mathcal{I}[B_1]$ and $F_2^\sharp \cup H_2 \in \mathcal{I}[B_2]$.

Let us prove this by showing first that the families of subsets extending $F_1^\flat$ and $F_2^\sharp$ within $\mathcal{I}$ are large; and then proceed to show that this large volume implies the existence of two subsets, H_1 and H_2 as required.

Let us first name these two families of subsets extending $F_1^\flat$ and $F_2^\sharp$ within $\mathcal{I}$:

$$\mathcal{I}_1 = \left\{ F \in \mathcal{P}\left(R_1\right) \;\middle|\; (F_1^\flat \cup F) \in \mathcal{I}[B_1] \right\} \quad \text{and} \quad \mathcal{I}_2 = \left\{ F \in \mathcal{P}\left(R_2\right) \;\middle|\; (F_2^\sharp \cup F) \in \mathcal{I}[B_2] \right\}$$

and proceed to prove they are large:

Proposition 4.11

$$\mu_q^{R_1}(\mathcal{I}_1) > \frac{1}{2} \quad \textit{and} \quad \mu_q^{R_2}(\mathcal{I}_2) > \frac{1}{2}$$

Proof: Let us prove the first case; the second one is proven by a symmetric, but otherwise identical, argument. By definition of $\mathcal{CF}_{B_1} = [\mathcal{I}[B_1]]_{C_1}^{\frac{3}{4}}$, it is the case that

$$\Pr_{F \in \mu_q}\left[F \in \mathcal{I}[B_1] \;\middle|\; F \cap C_1 = F_1^\flat \right] > \frac{3}{4}\,.$$

Note that the only difference between this event and

$$\mu_q^{R_1}(\mathcal{I}_1) = \Pr_{F \in \mu_q}\left[F \in \mathcal{I}[B_1] \;\middle|\; F \cap D_1 = F_1^\flat \right]$$

is the conditioning on F to not contain any block-assignment in $D_1 \setminus C_1$. Simplistically, if the elements in $D_1 \setminus C_1$ have tiny influence, then removing them from a subset does not take it out of $\mathcal{I}[B_1]$. Hence, it suffices to prove that this family, of extensions of $F_1^\flat$ within $\mathcal{I}[B_1]$, is almost independent of the set of block-assignments $D_1 \setminus C_1$, that is, that one can extract a small ($< \frac{1}{4}$) fraction of $\mathcal{I}_1$ and make it completely independent of the block-assignments outside $R_1 \cup C_1$.

Let us first observe that block-assignments in $D_1 \setminus C_1$ indeed have tiny influence,

Proposition 4.12

$$(D_1 \setminus C_1) \cap E_1 = \phi$$

Proof: There are two cases to consider for $\mathsf{a} \in D_1 \setminus C_1$: Either $\mathsf{a}|_{\hat{B}} \in \hat{C}_1$ and in that case, since $\hat{B}$ preserves B_1 and since $\mathsf{a} \notin C_1$, we deduce $\mathsf{a} \notin E_1$; or, $\mathsf{a}|_{\hat{B}} \in \hat{C}_2 \setminus \hat{C}_1$ and since the first condition on B_1 and B_2 in Proposition 4.7 is that $\hat{C}_1 \cap \mathbf{\Delta} = \hat{C}_2 \cap \mathbf{\Delta}$, we deduce $\mathsf{a}|_{\hat{B}} \notin \mathbf{\Delta}$. Now $\mathsf{a}|_{\hat{B}} \in \hat{C}_2 \subseteq \hat{E}_2$, implies $\mathsf{a}|_{\hat{B}} \notin \hat{E}_1$, thus $\mathsf{a} \notin E_1$. ■

By definition of the extended core E_i (Definition 4.1), it follows that for every $\mathsf{a} \in D_1 \setminus C_1$, $\mathsf{influence}_{\mathsf{a}}^q(\mathcal{I}[B_1]) < \eta$. Since $|D_1 \setminus C_1| < 4h_0$ (Proposition 4.10) we can deduce that $\mathcal{I}[B_1]$ is almost independent of $D_1 \setminus C_1$, utilizing a relatively simple, general property related to influences. Namely, that, given any monotone family of subsets of a domain R, and a set $U \subset R$ of elements of tiny influence, one has to remove only a small fraction of the family to make it completely independent of U, i.e. determined by $R \setminus U$. More accurately, we prove the following simple proposition in Appendix C,

Proposition 4.13 *Let $\mathcal{F} \subset \mathcal{P}(R)$ be monotone, and let $U \subset R$ be such that for all $e \in U$, $\mathsf{influence}_e^p(\mathcal{F}) < \eta$. Let*

$$\mathcal{F}' = \{ F \in \mathcal{F} \mid F \setminus U \in \mathcal{F} \}$$

then,

$$\mu_p^R\left(\mathcal{F} \setminus \mathcal{F}'\right) < |U| \cdot \eta \cdot p^{-|U|}$$

Proof: See Appendix C. ■

Substituting $D_1 \setminus C_1$ for U and $\frac{1}{16h_0} \cdot p^{5h_0}$ for η (see Definition 2.3), this proposition asserts that the weight of the subsets that have to be removed from $\mathcal{I}[B_1]$ to make it independent of $D_1 \setminus C_1$,

$$\mathcal{I}[B_1]' \stackrel{def}{=} \{F \in \mathcal{I}[B_1] \mid (F \setminus (D_1 \setminus C_1)) \notin \mathcal{I}[B_1]\},$$

is bounded by

$$\mu_q^{R_{B_1}}(\mathcal{I}[B_1]') < 4h_0 \cdot \eta \cdot q^{-4h_0} \leq \frac{1}{4} q^{h_0}.$$

Now, even if all $\mathcal{I}[B_1]'$ is concentrated on $F_1^\flat$, since $F_1^\flat$'s weight in $\mathcal{P}(C_1)$ is at least $q^{|C_1|} \geq q^{h_0}$, $\mu_q^{C_1}\left(F_1^\flat\right) \geq q^{h_0}$. It follows that (using $\Pr(A \mid B) \leq \Pr(A)/\Pr(B)$),

$$\Pr_{F \in \mu_q^{R_1}}\left[F \in \mathcal{I}[B_1]' \,\middle|\, F \cap C_1 = F_1^\flat\right] \leq \Pr_{F \in \mu_q^{R_1}}\left[F \in \mathcal{I}[B_1]'\right] \cdot \frac{1}{\mu_q^{C_1}\left(F_1^\flat\right)} < \frac{1}{4}$$

Formally, we write

$$\begin{aligned}
\frac{3}{4} &< \Pr\left[F \in \mathcal{I}[B_1] \mid F \cap C_1 = F_1^\flat\right] = \\
&= \Pr\left[F \in \mathcal{I}[B_1] \setminus \mathcal{I}[B_1]' \,\middle|\, F \cap C_1 = F_1^\flat\right] + \Pr\left[F \in \mathcal{I}[B_1]' \,\middle|\, F \cap C_1 = F_1^\flat\right] \\
&< \Pr\left[F \in \mathcal{I}[B_1] \setminus \mathcal{I}[B_1]' \,\middle|\, F \cap D_1 = F_1^\flat\right] + \frac{1}{4},
\end{aligned}$$

implying that $\mu_q^{R_1}(\mathcal{I}_1) = \Pr\left[F \in \mathcal{I}[B_1] \mid F \cap D_1 = F_1^\flat\right] > \frac{1}{2}$, and completing the proof of Proposition 4.11. ■

We complete the proof of the Soundness Lemma, by deducing from the large volume of I_1, I_2, the existence of two subsets $H_1 \in I_1$ and $H_2 \in I_2$ so that $\langle H_1, H_2 \rangle \in E_{\mathcal{B}}^{\mathbb{C}}$, implying $\left\langle F_1^\flat \cup H_1, F_2^\sharp \cup H_2 \right\rangle \in E_{\mathcal{B}}^{\mathbb{C}}$, which is the desired contradiction.

Proposition 4.14 *Let $I_1 \subset \mathcal{P}(R_1), I_2 \subset \mathcal{P}(R_2)$. If $(1-q)^2 \geq q$ and $\mu_q^{R_1}(\mathcal{I}_1) + \mu_q^{R_2}(\mathcal{I}_2) > 1$, there exist $H_1 \in \mathcal{I}_1$ and $H_2 \in \mathcal{I}_2$ such that $\langle H_1, H_2 \rangle \in E_{\mathcal{B}}^{\mathbb{C}}$.*

Proof: This proposition is proven by modifying the proof for the case of cross-intersecting families (Proposition 3.5). In that proof, we bounded the size of a pair of cross-intersecting families by pairing each subset with its complement, noting that at $p = \frac{1}{2}$ their weights are equal.

In this case, we focus on the value $q = p_{\max} = \frac{3-\sqrt{5}}{2}$ for which $(1-q)^2 = q$, noting that since $q \leq p_{\max}$, the monotonicity of I_1, I_2 (see Proposition 3.3) yields $\mu_{p_{\max}}(I_1) + \mu_{p_{\max}}(I_2) > 1$. Here let us partition both $\mathcal{P}(R_1)$ and $\mathcal{P}(R_2)$, and define an appropriate 'complement' differently for each part.

Fix an $(l-1)$-block assignment $\hat{\mathsf{a}} \in \hat{R}$. Extending it to a block assignment in R_1 (resp. in R_2) amounts to assigning a T/F value to v_1 (resp. to v_2). We denote these assignments by $\hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{T})}, \hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{F})} \in R_1$ and respectively $\hat{\mathsf{a}}_{(v_2 \leftarrow \mathsf{T})}, \hat{\mathsf{a}}_{(v_2 \leftarrow \mathsf{F})} \in R_2$. Our partition is defined according to a 'representative mapping' mapping each $F_1 \in \mathcal{P}(R_1)$ to a function $\Pi[F_1] : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$

defined as follows:

$$\forall \hat{\mathsf{a}} \in \hat{R}, \quad \Pi[F_1](\hat{\mathsf{a}}) \stackrel{def}{=} \begin{cases} \overline{\mathsf{TF}} & \hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{T})}, \hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{F})} \notin F_1 \\ \mathsf{T}\overline{\mathsf{F}} & \hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{T})} \in F_1,\ \hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{F})} \notin F_1 \\ \mathsf{F} & \hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{F})} \in F_1 \end{cases}$$

(symmetrically, we define $\Pi[F_2]$ for each $F_2 \in \mathcal{P}(R_2)$). This mapping is natural when considering the characteristic function of F_1 and asking, for every $\hat{\mathsf{a}} \in \hat{R}$, the value of that function on the two extensions of $\hat{\mathsf{a}}$ in R_1, $\hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{T})}$ and $\hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{F})}$.

Additionally, for a function $\Pi = \Pi[F_1]$, $\Pi : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$, let its complement be $\Pi^{\mathsf{c}} : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$ defined as follows:

$$\forall \hat{\mathsf{a}} \in \hat{R}, \quad \Pi^{\mathsf{c}}(\hat{\mathsf{a}}) \stackrel{def}{=} \begin{cases} \overline{\mathsf{TF}} & \Pi(\hat{\mathsf{a}}) = \mathsf{F} \\ \mathsf{T}\overline{\mathsf{F}} & \Pi(\hat{\mathsf{a}}) = \mathsf{T}\overline{\mathsf{F}} \\ \mathsf{F} & \Pi(\hat{\mathsf{a}}) = \overline{\mathsf{TF}}\,. \end{cases}$$

Observe that $\Pi^{\mathsf{cc}} = \Pi$, so that this is indeed a perfect matching of the possible functions $\Pi : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$. Most importantly, observe next that $\Pi[H_1] = \Pi^{\mathsf{c}}[H_2]$ implies $\langle H_1, H_2 \rangle \in E^{\mathbb{C}}_{\mathcal{B}}$. To see that, we need to verify that $H_1 \times H_2 \subset E_{\mathcal{B}}$. Indeed for every $\mathsf{a}_1 \in H_1, \mathsf{a}_2 \in H_2$, if $\mathsf{a}_1|_{\hat{B}} \neq \mathsf{a}_2|_{\hat{B}}$ then immediately $\langle \mathsf{a}_1, \mathsf{a}_2 \rangle \in E_{\mathcal{B}}$. More interestingly, if $\mathsf{a}_1|_{\hat{B}} = \mathsf{a}_2|_{\hat{B}} = \hat{\mathsf{a}}$ then it must be that $\Pi[H_1](\hat{\mathsf{a}}) = \mathsf{T}\overline{\mathsf{F}} = \Pi^{\mathsf{c}}[H_2](\hat{\mathsf{a}})$, namely $\mathsf{a}_1 = \hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{T})}$ and $\mathsf{a}_2 = \hat{\mathsf{a}}_{(v_2 \leftarrow \mathsf{T})}$. This again implies $\langle \mathsf{a}_1, \mathsf{a}_2 \rangle \in E_{\mathcal{B}}$ because by our assumption $\langle v_1, v_2 \rangle$ is an edge in G (part of an h-clique).

Next, observe that for a fixed $\Pi_0 : \hat{R} \to \left\{\overline{\mathsf{TF}}, \mathsf{T}\overline{\mathsf{F}}, \mathsf{F}\right\}$,

$$\Pr_{F_1 \in \mu_q^{R_1}} [\Pi[F_1] = \Pi_0] = \prod_{\hat{\mathsf{a}}:\, \Pi_0(\hat{\mathsf{a}}) = \overline{\mathsf{TF}}} (1-q)^2 \cdot \prod_{\hat{\mathsf{a}}:\, \Pi_0(\hat{\mathsf{a}}) = \mathsf{T}\overline{\mathsf{F}}} q(1-q) \cdot \prod_{\hat{\mathsf{a}}:\, \Pi_0(\hat{\mathsf{a}}) = \mathsf{F}} q\,.$$

Now if $q = p_{\max}$, i.e. $(1-q)^2 = q$, we have $\Pr_F[\Pi[F] = \Pi_0] = \Pr_F[\Pi[F] = \Pi_0^{\mathsf{c}}]$. Since $\mu_q(I_1) + \mu_q(I_2) > 1$, there must be a pair Π, Π^{c} such that

$$\{F_1 \in \mathcal{P}(R_1) \mid \Pi[F_1] = \Pi\} \cap I_1 \neq \phi \quad \text{and} \quad \{F_2 \in \mathcal{P}(R_2) \mid \Pi[F_2] = \Pi^{\mathsf{c}}\} \cap I_2 \neq \phi$$

providing the necessary pair of $H_1 = F_1^{\flat} \cup F_1 \in \mathcal{I}_1, H_2 = F_2^{\sharp} \cup F_2 \in \mathcal{I}_2$ with $\langle H_1, H_2 \rangle \in E^{\mathbb{C}}_{\mathcal{B}}$. ■

Lemma 4.6 is thereby proved. ■

This completes the proof of the soundness of the construction (Lemma 4.1). ■

The main theorem (Theorem 2.2) is thereby proven as well.

5 Tightness

In this section we show our analysis of $G^{\mathbb{C}}_{\mathcal{B}}$ is tight in two respects. First, we show the 2-intersecting bound: Namely, that for any value of p there is always an independent set $\mathcal{I}$ in $G^{\mathbb{C}}_{\mathcal{B}}$ whose size is almost $p^{\bullet}$, regardless of G being a 'yes' or a 'no' instance. Next, we show that if $p > (1-p)^2$ (this happens for $p > \frac{3-\sqrt{5}}{2}$), then a large independent set can be formed in $G^{\mathbb{C}}_{\mathcal{B}}$, again, regardless of the size of $\mathrm{IS}(G)$.

The 2-intersecting bound. We will exhibit an appropriate choice of maximal 2-intersecting families for almost all of the blocks $\mathcal{B}$, that constitutes an independent set in $G_{\mathcal{B}}^{\mathbb{C}}$.

According the the complete intersection theorem, when $p \approx \frac{3-\sqrt{5}}{2}$, the μ_p-largest 2-intersecting family is obtained by fixing some 4 block-assignments and taking all subsets that contain at least 3 of them. We will fix 4 block-assignments for almost all blocks. This will be done so that for every pair of these blocks, always at least 3 of the 4 block-assignments are pairwise consistent. Having a "3 out of 4" family of subsets by fixing these four elements will be an independent set.

Let $V_{red} \cup V_{green} \cup V_{blue} \cup V_{yellow}$ be an arbitrary partition of V, with roughly $|V|/4$ vertices in each. For every block $B \in \mathcal{B}$, define four special block-assignments, $\mathsf{a}^B_{red}, \mathsf{a}^B_{green}, \mathsf{a}^B_{blue}, \mathsf{a}^B_{yellow}$ defined as being true on their color, and false elsewhere, e.g.

$$\forall v \in B, \quad \mathsf{a}^B_{red}(v) \stackrel{def}{=} \begin{cases} \mathsf{T} & v \in V_{red} \cap B \\ \mathsf{F} & v \in B \setminus V_{red}\,. \end{cases}$$

Of course, not all four are well-defined for every block, as a block-assignment $\mathsf{a} \in R_B$ must contain at least t T's, and there is a negligible fraction of the blocks $\mathcal{B}' \subset \mathcal{B}$ that intersect at least one of $V_{red} \cup V_{green} \cup V_{blue} \cup V_{yellow}$ with less than t values. Neglecting these, we take for each block, the following set of vertices

$$\mathcal{I}[B] = \left\{ F \in V[B] \;\middle|\; \left| F \cap \left\{ \mathsf{a}^B_{red}, \mathsf{a}^B_{green}, \mathsf{a}^B_{blue}, \mathsf{a}^B_{yellow} \right\} \right| \geq 3 \right\}.$$

and let $\mathcal{I} \stackrel{def}{=} \bigcup_{B \in \mathcal{B} \setminus \mathcal{B}'} \mathcal{I}[B]$.

Let $\hat{B} \in V^{(l-1)}$, and let $B_1 = \hat{B} \cup \{v_1\}$, and $B_2 = \hat{B} \cup \{v_2\}$. Assume $v_1 \in V_{red}$ (symmetrically for any other color), and observe the following,

1. There is no edge in $E_{\mathcal{B}}$ between $\mathsf{a}^{B_1}_{green}$ and $\mathsf{a}^{B_2}_{green}$. Similarly, $\left\langle \mathsf{a}^{B_1}_{blue}, \mathsf{a}^{B_2}_{blue} \right\rangle, \left\langle \mathsf{a}^{B_1}_{yellow}, \mathsf{a}^{B_2}_{yellow} \right\rangle \notin E_{\mathcal{B}}$.

2. For any $F_1 \in \mathcal{I}[B_1]$, $\left| F_1 \cap \left\{ \mathsf{a}^{B_1}_{green}, \mathsf{a}^{B_1}_{blue}, \mathsf{a}^{B_1}_{yellow} \right\} \right| \geq 2$, and similarly for $F_2 \in \mathcal{I}[B_2]$. If wlog

   $$F_1 \cap \left\{ \mathsf{a}^{B_1}_{green}, \mathsf{a}^{B_1}_{blue}, \mathsf{a}^{B_1}_{yellow} \right\} = \left\{ \mathsf{a}^{B_1}_{green}, \mathsf{a}^{B_1}_{yellow} \right\} \quad and \quad F_2 \cap \left\{ \mathsf{a}^{B_2}_{green}, \mathsf{a}^{B_2}_{blue}, \mathsf{a}^{B_2}_{yellow} \right\} = \left\{ \mathsf{a}^{B_2}_{green}, \mathsf{a}^{B_2}_{blue} \right\},$$

   then F_1 is consistent with F_2 because of there being no edge in $E_{\mathcal{B}}$ between $\mathsf{a}^{B_1}_{green}$ and $\mathsf{a}^{B_2}_{green}$.

Thus, $\mathcal{I}$ is an independent set.

The bound $p \leq (1-p)^2$. Assume $p > \frac{3-\sqrt{5}}{2}$. We construct an independent set by selecting an arbitrary block assignment for each block, and taking all subsets containing it. By removing a negligible fraction of the vertices (subsets) in each block, we eliminate all edges between blocks.

Consider two blocks $B_1, B_2 \in \mathcal{B}$, such that $B_1 = \hat{B} \cup \{v_1\}, B_2 = \hat{B} \cup \{v_2\}$. Denote by $\hat{R}$ the set of sub-block assignments for $\hat{B}$ that are restrictions of R_{B_1} and of R_{B_2}, and assume for simplicity that every sub-block assignment in $\hat{R}$ has two extensions (to F and to T) in both R_{B_1} and R_{B_2}.

A random subset $F \in_{\mu_p} \mathcal{P}(R_{B_1})$, has expectedly $p \cdot |R_{B_1}|$ block-assignments. Moreover, there are expectedly $(1-p)^2 \cdot |\hat{R}|$ sub-block-assignments in $\hat{R}$ for which $\hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{F})}, \hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{T})} \notin F$, and expectedly $p \cdot |\hat{R}|$ sub-block-assignments for which $\hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{F})} \in F$.

For two vertices $F_1 \in V[B_1]$ and $F_2 \in V[B_2]$ to be inconsistent, one of them must deviate from the expectation, due to the following. Every $\hat{\mathsf{a}} \in \hat{R}$ for which $\hat{\mathsf{a}}_{(v_1 \leftarrow \mathsf{F})} \in F_1$ must have both $\hat{\mathsf{a}}_{(v_2 \leftarrow \mathsf{F})}, \hat{\mathsf{a}}_{(v_2 \leftarrow \mathsf{T})} \notin F_2$. If both F_1, F_2 are near their expectation, there are roughly $(1-p)^2 \cdot |\hat{R}|$ sub-block-assignments in $\hat{R}$ for which $\mathsf{a}_{(v_2 \leftarrow \mathsf{F})}, \mathsf{a}_{(v_2 \leftarrow \mathsf{T})} \notin F_2$, and if $(1-p)^2 < p$, this is not enough to meet the expected $p \cdot |\hat{R}|$ sub-block-assignments for which $\mathsf{a}_{(v_2 \leftarrow \mathsf{F})} \in F_1$.

Standard Chernoff bounds imply that we need to remove only a tiny fraction of the vertices of each block, so as to eliminate all subsets that deviate from the expectation according to at least one sub-block $\hat{B}$.

6 Discussion

Clearly, the most important open question we leave is finding the precise factor to within which the minimum vertex cover can be approximated. The results presented herein appear as partial progress towards resolving that question.

One of the more likely approaches would be to strengthen the structural properties of the graph $G_{\mathcal{B}}$, on which the biased Long-code is applied (replacing each block by the biased intersection graph).

Following our work, Khot [Kho02] has formulated a specific type of "unique-games" PCP system that implies such a structural restriction. Roughly, the property is that the constraints are on pairs of variables, and are bijective (for a constraint to be satisfied, every value for one variable leaves one value for the other, and vice versa). In that framework our graph $G_{\mathcal{B}}$ is equivalent to two-to-two constraints (where a value for one variable leaves at most two possible values). Khot raised the question of whether an NP-hardness result can still hold with such restricted constraints. In particular it was later shown in [KR03], that a construction as hinted above, but starting from Khot's *conjectured* "unique-games" PCP system, will establish an optimal hardness factor of $2-\epsilon$ for Minimum Vertex Cover, utilizing techniques presented herein.

Let us note that our result also implies, by direct reduction [Aro, Tre], a hardness of approximation of $(1.36)^2 \approx 1.84$ for the 2-CNF clause deletion problem: the problem of finding the minimum weight set of clauses in a 2-CNF formula, whose deletion makes the formula satisfiable. The best approximation algorithm for this problem guarantees only a factor of $\log n \log \log n$ [KPRT97].

The framework for proving hardness results suggested herein can be tried on other problems for which the known hardness result does not match the best upper-bound. Of these, particularly interesting are the problem of coloring 3-colorable graphs with fewest possible colors, and the problem of approximating the largest cut in a graph.

7 Acknowledgements

We would like to thank Noga Alon for his most illuminating combinatorial guidance, and Gil Kalai and Ehud Freidgut for highly influential discussions. We also thank Guy Kindler and Amnon Ta-Shma and Nati Linial for many constructive remarks on earlier versions of this paper, and Oded Goldreich for inspiring the presentation in its current more combinatorial form. In addition, we would like to thank the brave reading group who had helped us in constructive comments and to convince ourselves of the correctness of the proof: Oded Regev, Robi Krauthgamer, Vera Asodi, Oded Schwartz, Michael Langberg, Dana Moshkovich, Adi Akavia, and Elad Hazan. Thanks also to Sanjeev Arora for pointing out to us the application to the 2CNF deletion problem.

References

[ABSS97] S. Arora, L. Babai, J. Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes and linear equations. *J. Comput. Syst. Sci.*, 54(2):317–331, 1997.

[Ajt98] Miklós Ajtai. The shortest vector problem in L_2 is NP-hard for randomized reductions. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC-98)*, pages 10–19, New York, May 23–26 1998. ACM Press.

[AK97] Rudolf Ahlswede and Levon H. Khachatrian. The complete intersection theorem for systems of finite sets. *European J. Combin.*, 18(2):125–136, 1997.

[ALM$^+$98] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and intractability of approximation problems. *J. ACM*, 45(3):501–555, 1998.

[Aro] S. Arora. Personal communication.

[AS98] S. Arora and S. Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998.

[BGLR93] M. Bellare, S. Goldwasser, C. Lund, and A. Russell. Efficient multi-prover interactive proofs with applications to approximation problems. In *Proc. 25th ACM Symp. on Theory of Computing*, pages 113–131, 1993.

[BGS98] Mihir Bellare, Oded Goldreich, and Madhu Sudan. Free bits, PCPs, and nonapproximability—towards tight results. *SIAM Journal on Computing*, 27(3):804–915, June 1998.

[BK97] J. Bourgain and G. Kalai. Influences of variables and threshold intervals under group symmetries. *GAFA*, 7:438–461, 1997.

[BKS99] I. Benjamini, G. Kalai, and O. Schramm. Noise sensitivity of boolean functions and applications to percolation. *I.H.E.S.*, 90:5–43, 1999.

[BOL89] M. Ben Or and N. Linial. Collective coin flipping. In S. Micali, editor, *Randomness and Computation*, pages 91–115. Academic Press, 1989.

[BYE85] R. Bar-Yehuda and S. Even. A local-ratio theorem for approximating the weighted vertex cover problem. *Annals of Discrete Mathematics*, 25:27–45, 1985.

[Che52] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statistics*, 23:493–507, 1952.

[Coo71] S. Cook. The complexity of theorem-proving procedures. In *Proc. 3rd ACM Symp. on Theory of Computing*, pages 151–158, 1971.

[DGKR03] I. Dinur, V. Guruswami, S. Khot, and O. Regev. A new multilayered PCP and the hardness of hypergraph vertex cover. In *Proceedings of the 35th ACM Symposium on Theory of Computing (STOC)*, pages 595–601, 2003.

[DKRS03] I. Dinur, G. Kindler, R. Raz, and S. Safra. Approximating-CVP to within almost-polynomial factors is NP-hard. *Combinatorica*, 23(2):205–243, 2003.

[DRS02] Irit Dinur, Oded Regev, and Clifford D. Smyth. The hardness of 3-Uniform hypergraph coloring. In *Proceedings of the 43rd Symposium on Foundations of Computer Science (FOCS)*, pages 33–42, 2002.

[EKR61] P. Erdős, Chao Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.

[ER60] P. Erdős and R. Rado. Intersection theorems for systems of sets. *J. London Math. Soc.*, 35:85–90, 1960.

[Fei98] U. Feige. A threshold of ln n for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

[FF91] Peter Frankl and Zoltán Füredi. Beyond the Erdős-Ko-Rado theorem. *J. Combin. Theory Ser. A*, 56(2):182–194, 1991.

[FGL$^+$96] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Approximating clique is almost NP-complete. *Journal of the ACM*, 43(2):268–292, 1996.

[FK96] E. Friedgut and G. Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 124:2993–3002, 1996.

[FK98] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57(2):187–199, 1998.

[Fra78] P. Frankl. The Erdős-Ko-Rado theorem is true for $n = ckt$. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. I*, pages 365–375. North-Holland, Amsterdam, 1978.

[Fri98] Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998.

[Hal02] Eran Halperin. Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs. *Siam Journal on Computing*, 31:1608–1623, 2002.

[Hås99] Johan Håstad. Clique is hard to approximate within n to the power $1 - \epsilon$. *Acta Mathematica*, 182(1):105–142, 1999.

[Hås01] Johan Håstad. Some optimal inapproximability results. *Journal of ACM*, 48:798–859, 2001.

[Kar72] R. M. Karp. *Reducibility Among Combinatorial Problems*, pages 85–103. Plenum Press, 1972.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 767–775. ACM Press, 2002.

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In IEEE, editor, *29th annual Symposium on Foundations of Computer Science, October 24–26, 1988, White Plains, New York*, pages 68–80. IEEE Computer Society Press, 1988.

[KPRT97] Philip N. Klein, Serge A. Plotkin, Satish Rao, and Eva Tardos. Approximation algorithms for Steiner and directed multicuts. *Journal of Algorithms*, 22(2):241–269, 1997.

[KR03] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\varepsilon$. In *Proc. of 18th IEEE Annual Conference on Computational Complexity (CCC)*, pages 379–386, 2003.

[Lev73] L. Levin. Universal'nyĭe perebornyĭe zadachi (universal search problems : in Russian). *Problemy Peredachi Informatsii*, 9(3):265–266, 1973.

[LLL82] A.K. Lenstra, H.W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:513–534, 1982.

[LY94] Carsten Lund and Mihalis Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41(5):960–981, 1994.

[Mar74] G. Margulis. Probabilistic characteristics of graphs with large connectivity (in russian). *Probl. Pered. Inform.*, 10(2):101–108, 1974.

[Mic] D. Micciancio. The shortest vector in a lattice is hard to approximate to within some constant. *SIAM Journal on Computing*, 30(6):2008–2035, 2000.

[MS83] B. Monien and E. Speckenmeyer. Some further approximation algorithms for the vertex cover problem. In G. Ausiello and M. Protasi, editors, *Proceedings of the 8th Colloquium on Trees in Algebra and Programming (CAAP'83)*, volume 159 of *LNCS*, pages 341–349, L'Aquila, Italy, March 1983. Springer.

[PY91] C. Papadimitriou and M. Yannakakis. Optimization, approximation and complexity classes. *Journal of Computer and System Sciences*, 43:425–440, 1991.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, June 1998.

[RS97] R. Raz and S. Safra. A sub-constant error-probability low-degree test, and a sub-constant error-probability PCP characterization of NP. In *Proc. 29th ACM Symp. on Theory of Computing*, pages 475–484, 1997.

[Rus82] L. Russo. An approximative zero-one law. *Zeit. Warsch. und Verwandte Gebiete*, 61:129–139, 1982.

[Tal94] M. Talagrand. On Russo's approximate 0-1 law. *The Annals of Probability*, 22(3):1576–1587, 1994.

[Tre] L. Trevisan. Personal communication.

[Wil84] Richard M. Wilson. The exact bound in the Erdős-Ko-Rado theorem. *Combinatorica*, 4(2-3):247–257, 1984.

A Weighted vs Unweighted

Given a graph $G = (V, E, \Lambda)$, we construct, for any precision parameter $\varrho > 0$, an unweighted graph $G_\varrho = (V_\varrho, E_\varrho)$ with $\left|\frac{\overline{\mathrm{IS}}(G_\varrho)}{|V_\varrho|} - \overline{\mathrm{IS}}(G)\right| \leq \varrho$, and whose size is polynomial in $|G|$ and $\frac{1}{\varrho}$.

Let $n = |V| \cdot \frac{1}{\varrho}$. We replace each $v \in V$ with $n_v = \lceil n \cdot \Lambda(v) \rfloor$ copies ($\lceil x \rfloor$ denotes the integer nearest x), and set

$$\begin{aligned} V_\varrho &\stackrel{def}{=} \{\langle v, i\rangle \mid v \in V,\ 1 \leq i \leq n_v\}\ , \\ E_\varrho &\stackrel{def}{=} \{\{\langle v_1, i_1\rangle, \langle v_2, i_2\rangle\} \mid \{v_1, v_2\} \in E,\ i_1 \in [n_{v_1}],\ i_2 \in [n_{v_2}]\}\ . \end{aligned}$$

If $C \subseteq V$ is a vertex cover for G, then $C_\varrho = \bigcup_{v \in C} \{v\} \times [n_v]$ is a vertex cover for G_ϱ. Moreover, every minimal vertex cover $C_\varrho \subseteq V_\varrho$ is of this form, because whenever $\{v\} \times [n_v] \not\subseteq C_\varrho$ then by minimality $C_\varrho \cap (\{v\} \times [n_v]) = \phi$. Thus we show $\left|\frac{\overline{\mathrm{IS}}(G_\varrho)}{|V_\varrho|} - \overline{\mathrm{IS}}(G)\right| \leq \varrho$ by the following proposition,

Proposition A.1 *Let $C \subseteq V$, and let $C_\varrho = \bigcup_{v \in C} \{v\} \times [n_v]$. Then $\left|\frac{|C_\varrho|}{|V_\varrho|} - \Lambda(C)\right| \leq \varrho$.*

Proof: For every C, C_ϱ as above,

$$|C_\varrho| = \sum_{v \in C} n_v = \sum_{v \in C} \lceil n \cdot \Lambda(v) \rfloor = n \cdot \Lambda(C) + \sum_{v \in C} (\lceil n \cdot \Lambda(v) \rfloor - n \cdot \Lambda(v)).$$

For any v, $|\lceil v \rfloor - v| \leq \frac{1}{2}$, and so

$$\left|\frac{|C_\varrho|}{n} - \Lambda(C)\right| \leq \frac{1}{2}\frac{|C|}{n} \leq \frac{\varrho}{2} \tag{5}$$

To complete our proof we need to replace $\frac{|C_\varrho|}{n}$ by $\frac{|C_\varrho|}{|V_\varrho|}$ in (5). Indeed, taking $C = V$ in (5), yields $\left|\frac{|V_\varrho|}{n} - 1\right| \leq \frac{\varrho}{2}$, and multiplying by $\frac{|C_\varrho|}{|V_\varrho|} \leq 1$, we obtain $\left|\frac{|C_\varrho|}{n} - \frac{|C_\varrho|}{|V_\varrho|}\right| \leq \frac{\varrho}{2}$. ■

B Proof of Theorem 2.1

In this section we prove Theorem 2.1 which encapsulates our use of the PCP theorem. PCP characterizations of NP in general state that given some SAT instance, namely, a set of Boolean-functions $\Phi = \{\varphi_1, ..., \varphi_n\}$ over variables W, it is NP-hard to distinguish between 'yes' instances where there is an assignment A to Φ's variables that satisfies all Φ, and 'no' instances where any assignment to A satisfies at most a small fraction of Φ.

Definition B.1 *Denote by $\Upsilon(\Phi)$ the maximum, over all assignments to Φ's variables $A : W \to \{0, 1\}$, of the fraction of $\varphi \in \Phi$ satisfied by A, namely*

$$\Upsilon(\Phi) = \max_{A} \Pr_{\varphi \in \Phi}[\varphi \textit{ is satisfied by } A]$$

The basic PCP theorem showing hardness for gap-SAT states that,

Theorem B.1 ([AS98, ALM+98]) *There exists some constant $\beta > 0$ such that given a set $\Phi = \{\varphi_1, .., \varphi_n\}$ of 3-CNF clauses over Boolean variables W (each clause is the OR of exactly 3 variables), it is NP-hard to distinguish between the two cases:*

Yes: Φ *is satisfiable* ($\Upsilon(\Phi) = 1$).

No: $\Upsilon(\Phi) < 1 - \beta$.

■

Let us first sketch the proof for Theorem 1.5, based on the above Theorem and the Parallel-repetition theorem [Raz98], and then turn to the consequent proof of Theorem 2.1.

Proof: Given Φ as above, define the parallel repetition version of Φ:

Definition B.2 ($\mathsf{Par}\,[\Phi, k]$) *Let* $\langle \Phi, W \rangle$ *be a 3-CNF instance, with 3-CNF clauses* Φ *over variables* W. *For any integer* $k > 0$, *let*

$$\mathsf{Par}\,[\Phi, k] \stackrel{def}{=} \langle \Psi, X, Y \rangle$$

be a SAT instance with Boolean functions Ψ *over two types of variables:* $X \stackrel{def}{=} \Phi^k$ *and* $Y \stackrel{def}{=} W^k$.

The range of each variable $x = (\varphi_1, \ldots, \varphi_k) \in X$, *is* $R_X = [7]^k$, *corresponding (by enumerating the* 7 *satisfying assignments of each 3-CNF clause* $\varphi \in \Phi$*) to the concatenation of the satisfying assignments for* $\varphi_1, \ldots, \varphi_k$. *The range of each variable* $y = (w_1, \ldots, w_k) \in Y$, *is* $R_Y = [2]^k$, *corresponding to all possible assignments to* $w_1, \ldots, w_k$.

For $y = (w_1, .., w_k)$ *and* $x = (\varphi_1, .., \varphi_k)$, *denote* $y \sqsubseteq x$ *if for all* $i \in [k]$, w_i *is a variable of* φ_i. *The Boolean-functions in* Ψ *are as follows:*

$$\Psi = \left\{ \psi_{x \to y} \;\middle|\; y \in W^k,\ x \in \Phi^k,\ y \sqsubseteq x \right\}$$

where $\psi_{x \to y}$ *is* T *if the assignment to* y *is the restriction to* y *of the assignment to* x, *and* F *otherwise. Since each test* $\varphi \in \Phi$ *has exactly* 3 *variables, each variable* $x \in X$ *appears in exactly* 3^k *tests in* $\psi_{x \to y} \in \Psi$.

Clearly, if $\Upsilon(\Phi) = 1$, then $\Upsilon(\Psi) = 1$. Moreover,

Theorem B.2 (Parallel Repetition, [Raz98]) *There exists some constant* $c > 0$, *such that, let* $\langle \Phi, W \rangle$ *be a 3-CNF-instance, and let* $\langle \Psi, X, Y \rangle = \mathsf{Par}\,[\Phi, k]$,

$$\Upsilon(\Psi) \leq \Upsilon(\Phi)^{c \cdot k}.$$

■

Therefore, one may choose k for which $(1 - \beta)^{c \cdot k} \leq \epsilon / h^3$ and $|R_Y|, |R_X| \leq (\frac{\epsilon}{h})^{-O(1)}$, hence it is NP-hard to distinguish whether $\Upsilon(\Psi) = 1$ or $\Upsilon(\Psi) < \epsilon / h^3$. ■

Now we may proceed to proving the following:

Theorem 2.1 *For any* $h, \epsilon > 0$, *the problem* $h\mathrm{IS}(r, \epsilon, h)$ *is NP-hard, as long as* $r \geq (\frac{h}{\epsilon})^c$ *for some constant* c.

Proof: By reduction from the above Theorem. Assume Ψ as above, and let us apply the FGLSS Construction [FGL+96, Kar72] to Ψ, specified next. Let $\mathcal{G}[\Psi]$ be the (m, r)-co-partite graph, with $m = |X|$ and $r = |R_X|$,

$$\mathcal{G}[\Psi] = \langle V, E \rangle \text{ where } V \stackrel{def}{=} (X \times R_X)$$

that is, where $\mathcal{G}[\Psi]$'s vertices is the set of pairs consisting of a variable x in X and a value $a \in R_X$ for x. For the edge set E of $\mathcal{G}[\Psi]$, let us consider all pairs of vertices whose values cannot possibly correspond to the same satisfying assignment:

$$E = \{\{(x_1,a_1),(x_2,a_2)\} \mid \exists y, \psi_{x_1\to y}, \psi_{x_2\to y} \in \Phi,\ \psi_{x_1\to y}(a_1) \neq \psi_{x_2\to y}(a_2)\}$$

Therefore, an independent set in $\mathcal{G}[\Psi]$ cannot correspond to an inconsistent assignment to Φ.

If Ψ is satisfiable, let $A : X \cup Y \to \{\mathsf{T},\mathsf{F}\}$ be a satisfying assignment for it, and observe that the set $\{(x, A(x)) \mid x \in X\} \subset V$ is an independent set of size $|X| = m$.

Otherwise, let us assume a set of vertices $\mathcal{I} \subset V$ in $\mathcal{G}[\Psi]$ that contains no clique of size h, and such that $|\mathcal{I}| > \epsilon \cdot m$, and show that $\Upsilon(\Psi) > \frac{\epsilon}{h^3}$. Let $A_{\mathcal{I}}$ map to each variable a subset of its range, as follows. For every $x \in X$ and $y \in Y$, set

$$\begin{aligned} A_{\mathcal{I}}(x) &\stackrel{def}{=} \{a \in R_X \mid (x,a) \in \mathcal{I}\} \\ A_{\mathcal{I}}(y) &\stackrel{def}{=} \bigcup_{\psi_{x\to y}\in\Psi} \psi_{x\to y}(A_{\mathcal{I}}(x)) \,. \end{aligned}$$

The key is that the h-clique freeness implies that for every $x \in X$, $|A_{\mathcal{I}}(x)| < h$ and for every $y \in Y$, $|A_{\mathcal{I}}(y)| < h$. Otherwise, if $|A_{\mathcal{I}}(y)| \geq h$ for some y, there are vertices $(x_1,a_1),\ldots,(x_h,a_h)$ so that $\{\psi_{x_i\to y}(a_i)\}$ are distinct hence these vertices form a clique of size h. By the definition of $A_{\mathcal{I}}$, for every x with $A_{\mathcal{I}}(x) \neq \phi$ and for *every* $\psi_{x\to y} \in \Psi$,

$$\psi_{x\to y}(A_{\mathcal{I}}(x)) \cap A_{\mathcal{I}}(y) \neq \phi \,.$$

Denote $X_0 = \{x \in X \mid A_{\mathcal{I}}(x) \neq \phi\}$ and observe that since there is an equal number of $\psi_{x\to y} \in \Psi$ for each variable x:

$$\Pr_{\psi_{x\to y}\in\Psi}[\psi_{x\to y}(A_{\mathcal{I}}(x)) \cap A_{\mathcal{I}}(y) \neq \phi] = \Pr_{x\in X}[x \in X_0] = \frac{|X_0|}{|X|} > \frac{1}{h}\cdot\frac{|\mathcal{I}|}{|X|} > \epsilon/h \,.$$

Finally, by picking for each variable $x \in X, y \in Y$ a random assignment

$$\forall x \in X, y \in Y, \qquad a_x \in_R A_{\mathcal{I}}(x), \quad a_y \in_R A_{\mathcal{I}}(y)$$

If $A_{\mathcal{I}}(x) \neq \phi$, the probability that $\psi_{x\to y} \in \Psi$ is satisfied by such a random assignment is at least $\frac{1}{|A_{\mathcal{I}}(x)|}\cdot\frac{1}{|A_{\mathcal{I}}(y)|} > 1/h^2$. Thus the expected number of Boolean functions satisfied by this random assignment is $> \frac{\epsilon}{h^3}\cdot|\Psi|$. Since at least one assignment must meet the expectation, $\Upsilon(\Psi) > \frac{\epsilon}{h^3}$.
■

C Some Propositions about μ_p

Proposition 3.3 *For a monotone family of subsets $\mathcal{F} \subseteq \mathcal{P}(n)$, $q > p \Rightarrow \mu_q(\mathcal{F}) \geq \mu_p(\mathcal{F})$.*

Proof: For a subset $F \in \mathcal{P}([n])$ denote

$$F_{\leq i} \stackrel{def}{=} F \cap [1,i] \quad \text{and} \quad F_{>i} \stackrel{def}{=} F \cap [i+1,n]$$

and consider, for $0 \leq i \leq n$, the hybrid distribution, where the first i elements are chosen with bias p and the others are chosen with bias q

$$\mu_{p,i,q}(F) \stackrel{def}{=} p^{|F_{\leq i}|}\cdot(1-p)^{i-|F_{\leq i}|}\cdot q^{|F_{>i}|}\cdot(1-q)^{n-i-|F_{>i}|}$$

Observe that

$$\forall 0 \le i \le n \quad \mu_{p,i,q}(\mathcal{F}) \ge \mu_{p,i+1,q}(\mathcal{F})$$

therefore $\mu_q(\mathcal{F}) = \mu_{p,0,q}(\mathcal{F}) \ge \mu_{p,n,q} = \mu_p(\mathcal{F})$. ■

Theorem 3.4 [Russo-Margulis Identity] *Let $\mathcal{F} \subseteq \mathcal{P}(R)$ be a monotone family. Then,*

$$\frac{d\mu_q(\mathcal{F})}{dq} = \mathsf{as}_q(\mathcal{F})$$

Proof: For a subset $F \in \mathcal{P}(n)$ write

$$\mu_q(F) = \prod_{i \in [n]} \mu_q^i(F), \qquad \text{for} \quad \mu_q^i(F) = \begin{cases} q & i \in F \\ \\ 1-q & i \notin F \end{cases} \tag{6}$$

Observe that

$$\mathsf{influence}_i^q(\mathcal{F}) = \sum_{F \in \mathcal{F}} \left(\frac{d\mu_q^i(F)}{dq} \cdot \prod_{j \ne i} \mu_q^j(F) \right)$$

Differentiating (6) according to q, and summing over all $F \in \mathcal{F}$, we get

$$\frac{d\mu_q(\mathcal{F})}{dq} = \sum_{i \in [n]} \mathsf{influence}_i^q(\mathcal{F}) = \mathsf{as}_q(\mathcal{F})\,.$$

■

We next show that for any monotone family $\mathcal{F} \subset \mathcal{P}(R)$, if $U \subset R$ is a set of elements of tiny influence, one has to remove only a small fraction of $\mathcal{F}$ to make it completely independent of U:

Proposition 4.13 *Let $\mathcal{F} \subset \mathcal{P}(R)$ be monotone, and let $U \subset R$ be such that for all $e \in U$, $\mathsf{influence}_e^p(\mathcal{F}) < \eta$. Let*

$$\mathcal{F}' = \{ F \in \mathcal{F} \mid F \setminus U \in \mathcal{F} \}$$

then,

$$\mu_p^R\left(\mathcal{F} \setminus \mathcal{F}'\right) < |U| \cdot \eta \cdot p^{-|U|}$$

Proof: Let

$$\mathcal{F}'' = \{ F \in \mathcal{P}(R \setminus U) \;\mid\; F \cup U \in \mathcal{F} \text{ but } F \notin \mathcal{F} \}\,.$$

A set $F \in \mathcal{F}''$ contributes at least $\mu_p^{R \setminus U}(F) \cdot p^{|U|}$ to the influence of at least one element $e \in U$. Since the sum of influences of elements in U is $< |U| \cdot \eta$, we have $\mu_p^{R\setminus U}(\mathcal{F}'') < |U| \cdot \eta \cdot p^{-|U|}$. The proof is complete noting that,

$$\mathcal{F} \setminus \mathcal{F}' \;\subseteq\; \left\{ F \mid F \cap (R \setminus U) \in \mathcal{F}'' \right\}$$

■

D Erdős-Ko-Rado

In this section we prove a lemma that is a continuous version and follows directly from the complete intersection theorem of Ahlswede and Khachatrian [AK97].

Let us define

$$\mathcal{A}_i \overset{def}{=} \{F \in \mathcal{P}([n]) \mid F \cap [1, 2+2i] \geq 2+i\}$$

and prove the following lemma,

Lemma 1.4 *Let $\mathcal{F} \subset \mathcal{P}([n])$ be 2-intersecting. For any $p < \frac{1}{2}$,*

$$\mu_p(\mathcal{F}) \leq \max_i \{\mu_p(\mathcal{A}_i)\}.$$

Proof: Denote $\mu = \max_i(\mu_p(\mathcal{A}_i))$. Assuming $\mathcal{F}_0 \subset \mathcal{P}([n_0])$ contradicts the claim, let $a = \mu_p(\mathcal{F}_0) - \mu > 0$. Now consider $\mathcal{F} = \mathcal{F}_0 \sqcup \mathcal{P}([n] \setminus [n_0])$ for $n > n_0$ large enough, to be determined later. Clearly, for any $n \geq n_0$, $\mu_p^{[n]}(\mathcal{F}) = \mu_p^{[n_0]}(\mathcal{F}_0)$, and $\mathcal{F}$ is 2-intersecting. Consider, for $\theta < \frac{1}{2} - p$ to be determined later,

$$S \overset{def}{=} \{k \in \mathbb{N} \mid \ |k - p \cdot n| \leq \theta \cdot n\}$$

and for every $k \in S$, denote by $\mathcal{F}_k = \mathcal{F} \cap \binom{[n]}{k}$. We will show that since most of $\mathcal{F}$'s weight is derived from $\cup_{k \in S} \mathcal{F}_k$, there must be at least one $\mathcal{F}_k$ that contradicts Theorem 3.7. Indeed,

$$\mu + a = \mu_p(\mathcal{F}) = \sum_{k \in S} p^k (1-p)^{n-k} \cdot |\mathcal{F}_k| + o(1)$$

Hence there exists $k \in S$ for which $\frac{|\mathcal{F}_k|}{\binom{[n]}{k}} \geq \mu + \frac{1}{2}a$. We have left to see that $\mu \cdot \binom{n}{k}$ is close enough to $\max_i(|\mathcal{A}_i \cap \binom{[n]}{k}|)$. This follows from usual tail bounds, and is sketched as follows. Subsets in $\binom{[n]}{k}$ for large enough i (depending only on $\frac{k}{n}$ but not on k or n), have roughly $\frac{k}{n} \cdot (2i+2)$ elements in the set $[1, 2i+2]$. Moreover, the subsets in $\mathcal{A}_i$ have at least $i+2$ elements in $[1, 2i+2]$, thus are very few (compared to $\binom{n}{k}$), because $\frac{i+2}{2i+2} > \frac{1}{2} > p + \theta \geq \frac{k}{n}$. In other words, there exists some constant $C_{p+\theta,\mu}$, for which $\left|\mathcal{A}_i \cap \binom{[n]}{k}\right| < \mu \cdot \binom{n}{k}$ for all $i \geq C_{p,\mu}$ as long as $\frac{k}{n} \leq p + \theta$.

Additionally, for every $i < C_{p,\mu}$, taking n to be large enough we have

$$\forall k \in S, \quad \frac{\left|\mathcal{A}_i \cap \binom{[n]}{k}\right|}{\binom{n}{k}} = \mu_{\frac{k}{n}}(\mathcal{A}_i) + o(1) = \mu_p(\mathcal{A}_i) + o(1) < \mu + o(1)$$

where the first equality follows from a straightforward computation. ■

We have the following corollary,

Corollary 3.8 *Let $\mathcal{F} \subset \mathcal{P}(R)$ be 2-intersecting. For any $q < p_{\max}$, $\mu_q(\mathcal{F}) \leq q^{\bullet}$.*

Proof: Define a sequence $p_0 < p_1 < \ldots$, by $p_i \overset{def}{=} \frac{i}{2i+1}$. We show that these are the points where the maximum switches from $\mathcal{A}_i$ to $\mathcal{A}_{i+1}$. More accurately, we show for all $i \geq 0$,

$$\forall p \in (p_i, p_{i+1}] \quad \max_j \{\mu_p(\mathcal{A}_j)\} = \mu_p(\mathcal{A}_i) \tag{7}$$

This, together with Lemma 1.4, implies the corollary, as $p < p_{\max} < 0.4 = p_2$ implies $\mu_p(\mathcal{F}) \le \max(\mu_p(\mathcal{A}_0), \mu_p(\mathcal{A}_1)) = \max(p^2, 4p^3 - 3p^4) = p^\bullet$.

So we proceed to prove (7). A subset $F \notin \mathcal{A}_i$ must intersect $[1, 2i+2]$ on at most $i+1$ elements. If additionally $F \in \mathcal{A}_{i+1}$ it must then contain $2i+3, 2i+4$. Thus,

$$\mu_p(\mathcal{A}_{i+1} \setminus \mathcal{A}_i) = \binom{2i+2}{i+1} \cdot p^{i+1}(1-p)^{i+1} \cdot p^2$$

Similarly,

$$\mu_p(\mathcal{A}_i \setminus \mathcal{A}_{i+1}) = \binom{2i+2}{i+2} \cdot p^{i+2}(1-p)^{i} \cdot (1-p)^2$$

Together,

$$\begin{aligned} \mu_p(\mathcal{A}_{i+1}) - \mu_p(\mathcal{A}_i) &= \mu_p(\mathcal{A}_{i+1} \setminus \mathcal{A}_i) - \mu_p(\mathcal{A}_i \setminus \mathcal{A}_{i+1}) \\ &= p^{i+2}(1-p)^{i+1}\binom{2i+2}{i+1}\left(p - (1-p)\frac{i+1}{i+2}\right) \end{aligned}$$

The sign of this difference is determined by $p - (1-p)\frac{i+1}{i+2}$. For a fixed $i \ge 0$, this expression goes from positive to negative passing through zero once at $p = \frac{i+1}{2i+3} = p_{i+1}$. Thus, the sequence $\{\mu_p(\mathcal{A}_j)\}_j$ is maximized at i for $p_i < p \le p_{i+1}$. (It is increasing when $i \le \frac{1-3p}{2p-1}$, and decreasing thereafter). ■

E A Chernoff Bound

Proposition E.1 *For any set $I \subset V$ such that $|I| = \frac{1}{r} \cdot |V|$,*

$$\Pr_{B \in \mathcal{B}}\left[|I \cap B| < l_{\mathsf{T}}\right] < 2e^{-\frac{2l_{\mathsf{T}}}{8}}$$

Proof: Consider the random variable $\chi_I : V \to \{0,1\}$ taking a 1 iff $v \in I$. We have $\Pr_{v \in V}[\chi_I(v) = 1] = \frac{1}{r}$, and for every $B \in \mathcal{B} = \binom{V}{l}$, $|I \cap B| = \sum_{v \in B} \chi_I(v)$, so the expectation of this is $|B| \cdot \frac{1}{r} = 2l_{\mathsf{T}}$. The standard Chernoff bound [Che52] directly gives

$$\Pr_{v_1,..,v_l \in V}\left[\sum_{i \in [l]} \chi_I(v_i) < l_{\mathsf{T}} = \frac{1}{2} \cdot l/r\right] < e^{-\frac{l}{8r}}$$

We are almost done, except that the above probability was taken with repetitions, while in our case, for $v_1, \ldots, v_l$ to constitute a block $B \in \mathcal{B}$, they must be l distinct values. In fact, this happens with overwhelming probability and in particular $\ge \frac{1}{2}$, thus we write,

$$\begin{aligned} \Pr_{v_1,\ldots,v_l \in V}\left[\sum_i \chi_I(v_i) < l_{\mathsf{T}} \;\middle|\; |\{v_1, \ldots, v_l\}| = l\right] &\le \frac{\Pr_{v_1,\ldots,v_l \in V}\left[\sum_i \chi_I(v_i) < l_{\mathsf{T}}\right]}{\Pr_{v_1,\ldots,v_l \in V}\left[|\{v_1, \ldots, v_l\}| = l\right]} \\ &\le \frac{e^{-\frac{l}{8r}}}{\frac{1}{2}} = 2e^{-\frac{l}{8r}} \end{aligned}$$

■